

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
29 March 2001 (29.03.2001)

PCT

(10) International Publication Number
WO 01/21189 A1

(51) International Patent Classification⁷: **A61K 38/00**,
38/04, 38/08, 38/10, 39/29, 39/295

MN 55902 (US). **KUBO, Ralph, T.** [US/US]; 6921 Pear
Tree Drive, Carlsbad, CA 92009 (US). **GREY, Howard,**
M. [US/US]; 1461 Caminito Batea, La Jolla, CA 92037
(US).

(21) International Application Number: PCT/US00/19774

(22) International Filing Date: 19 July 2000 (19.07.2000) (74) Agents: **LOCKYER, Jean, M.** et al.; Townsend and
Townsend and Crew LLP, Two Embarcadero Center,
Eighth Floor, San Francisco, CA 94111-3834 (US).

(25) Filing Language: English

(26) Publication Language: English (81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(30) Priority Data:
09/357,737 19 July 1999 (19.07.1999) US

(71) Applicant (for all designated States except US): **EPIM-
MUNE INC.** [US/US]; 5820 Nancy Ridge Drive, San
Diego, CA 92121 (US). (84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SI, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **SETTE, Alessandro**
[IT/US]; 5551 Linda Rosa Avenue, La Jolla, CA 92037
(US). **SIDNEY, John** [US/US]; 4218 Corte de la Siena,
San Diego, CA 92130 (US). **SOUTHWOOD, Scott**
[US/US]; 10679 Strathmore Drive, Santee, CA 92071
(US). **LIVINGSTON, Brian, D.** [US/US]; 13555 Chaco
Court, San Diego, CA 92129 (US). **CHESNUT, Robert**
[US/US]; 1473 Kings Cross Drive, Cardiff-by-the-Sea,
CA 92007 (US). **BAKER, Denise, Marie** [US/US]; 11575
Caminito LaBar #21, San Diego, CA 92126 (US). **CELIS,
Esteban** [US/US]; 3683 Wright Road S.W., Rochester,

Published:

— With international search report.

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 01/21189 A1

(54) Title: INDUCING CELLULAR IMMUNE RESPONSES TO HEPATITIS C VIRUS USING PEPTIDE AND NUCLEIC ACID
COMPOSITIONS

(57) Abstract: This invention uses our knowledge of the mechanisms by which antigen is recognized by T cells to identify and prepare HCV epitopes, and to develop epitope-based vaccines directed towards HCV. More specifically, this application communicates our discovery of pharmaceutical compositions and methods of use in the prevention and treatment of HCV infection.

**INDUCING CELLULAR IMMUNE RESPONSES TO HEPATITIS C VIRUS
USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS**

FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

10 This invention was funded, in part, by the United States government under grants with the National Institutes of Health. The U.S. government has certain rights in this invention.

INDEX

- 15 I. Background of the Invention
- II. Summary of the Invention
- III. Brief Description of the Figures
- IV. Detailed Description of the Invention
 - A. Definitions
 - 20 B. Stimulation of CTL and HTL responses
 - C. Binding Affinity of Peptide Epitopes for HLA Molecules
 - D. Peptide Epitope Binding Motifs and Supermotifs
 - 1. HLA-A1 supermotif
 - 2. HLA-A2 supermotif
 - 25 3. HLA-A3 supermotif
 - 4. HLA-A24 supermotif
 - 5. HLA-B7 supermotif
 - 6. HLA-B27 supermotif
 - 7. HLA-B44 supermotif
 - 30 8. HLA-B58 supermotif
 - 9. HLA-B62 supermotif
 - 10. HLA-A1 motif

11. HLA-A2.1 motif
12. HLA-A3 motif
13. HLA-A11 motif
14. HLA-A24 motif
- 5 15. HLA-DR-1-4-7 supermotif
16. HLA-DR3 motifs
- E. Enhancing Population Coverage of the Vaccine
- F. Immune Response-Stimulating Peptide Epitope Analogs
- G. Computer Screening of Protein Sequences from Disease-Related Antigens
- 10 for Supermotif- or Motif-Containing Epitopes
- H. Preparation of Peptide Epitopes
- I. Assays to Detect T-Cell Responses
- J. Use of Peptide Epitopes for Evaluating Immune Responses
- K. Vaccine Compositions
- 15 1. Minigene Vaccines
2. Combinations of CTL Peptides with Helper Peptides
- L. Administration of Vaccines for Therapeutic or Prophylactic Purposes
- M. Kits
- V. Examples
- 20 VI. Claims
- VII. Abstract

I. BACKGROUND OF THE INVENTION

Hepatitis C virus (HCV) infection is a global human health problem with approximately 150,000 new reported cases each year in the U.S. alone. HCV is a single stranded RNA virus, and is the etiological agent identified in most cases of non-A, non-B post-transfusion and post-transplant hepatitis, and is a common cause of acute sporadic hepatitis (Choo *et al.*, *Science* 244:359, 1989; Kuo *et al.*, *Science* 244:362, 1989; and Alter *et al.*, in: *Current Perspective in Hepatology*, p. 83, 1989). It is estimated that more than 50% of patients infected with HCV become chronically infected and, of those, 20% develop cirrhosis of the liver within 20 years (Davis *et al.*, *New Engl. J. Med.* 321:1501,

1989; Alter *et al.*, in: *Current Perspective in Hepatology*, p. 83, 1989; Alter *et al.*, *New Engl. J. Med.* 327:1899, 1992; and Dienstag, J. L. *Gastroenterology* 85:430, 1983). Moreover, the only therapy available for treatment of HCV infection is interferon- α . Most patients are unresponsive, however, and among the responders, there is a high 5 recurrence rate within 6-12 months of cessation of treatment (Liang *et al.*, *J. Med. Virol.* 40:69, 1993). Ribaviron, a guanosine analog with a broad spectrum activity against many RNA and DNA viruses, has been shown in clinical trials to be effective against chronic HCV infection when used in combination with interferon- α (see, e.g., Poynard *et al.*, *Lancet* 352:1426-1432, 1998; Reichard *et al.*, *Lancet* 351:83-87, 1998) However, the 10 response rate is still well below 50%.

Virus-specific, human leukocyte antigen (HLA) class I-restricted cytotoxic T lymphocytes (CTL) are known to play a major role in the prevention and clearance of virus infections *in vivo* (Oldstone *et al.*, *Nature* 321:239, 1989; Jamieson *et al.*, *J. Virol.* 61:3930, 1987; Yap *et al.*, *Nature* 273:238, 1978; Lukacher *et al.*, *J. Exp. Med.* 160:814, 15 1994; McMichael *et al.*, *N. Engl. J. Med.* 309:13, 1983; Sethi *et al.*, *J. Gen. Virol.* 64:443, 1983; Watari *et al.*, *J. Exp. Med.* 165:459, 1987; Yasukawa *et al.*, *J. Immunol.* 143:2051, 1989; Tigges *et al.*, *J. Virol.* 66:1622, 1993; Reddenhase *et al.*, *J. Virol.* 55:263, 1985; Quinnan *et al.*, *N. Engl. J. Med.* 307:6, 1982). HLA class I molecules are expressed on the surface of almost all nucleated cells. Following intracellular processing of antigens, 20 epitopes from the antigens are presented as a complex with the HLA class I molecules on the surface of such cells. CTL recognize the peptide-HLA class I complex, which then results in the destruction of the cell bearing the HLA-peptide complex directly by the CTL and/or via the activation of non-destructive mechanisms e.g., the production of interferon, that inhibit viral replication.

25 In view of the heterogeneous immune response observed with HCV infection, induction of a multi-specific cellular immune response directed simultaneously against multiple HCV epitopes appears to be important for the development of an efficacious vaccine against HCV. There is a need, however, to establish vaccine embodiments that elicit immune responses that correspond to responses seen in patients that clear HCV 30 infection.

The information provided in this section is intended to disclose the presently understood state of the art as of the filing date of the present application. Information is included in this section which was generated subsequent to the priority date of this

application. Accordingly, information in this section is not intended, in any way, to delineate the priority date for the invention.

II. SUMMARY OF THE INVENTION

5 This invention applies our knowledge of the mechanisms by which antigen is recognized by T cells, for example, to develop epitope-based vaccines directed towards HCV. More specifically, this application communicates our discovery of specific epitope pharmaceutical compositions and methods of use in the prevention and treatment of HCV infection.

10 Upon development of appropriate technology, the use of epitope-based vaccines has several advantages over current vaccines, particularly when compared to the use of whole antigens in vaccine compositions. There is evidence that the immune response to whole antigens is directed largely toward variable regions of the antigen, allowing for immune escape due to mutations. The epitopes for inclusion in an epitope-based vaccine 15 are selected from conserved regions of viral or tumor-associated antigens, which thereby reduces the likelihood of escape mutants. Furthermore, immunosuppressive epitopes that may be present in whole antigens can be avoided with the use of epitope-based vaccines.

20 An additional advantage of an epitope-based vaccine approach is the ability to combine selected epitopes (CTL and HTL), and further, to modify the composition of the epitopes, achieving, for example, enhanced immunogenicity. Accordingly, the immune response can be modulated, as appropriate, for the target disease. Similar engineering of the response is not possible with traditional approaches.

25 Another major benefit of epitope-based immune-stimulating vaccines is their safety. The possible pathological side effects caused by infectious agents or whole protein antigens, which might have their own intrinsic biological activity, is eliminated.

30 An epitope-based vaccine also provides the ability to direct and focus an immune response to multiple selected antigens from the same pathogen. Thus, patient-by-patient variability in the immune response to a particular pathogen may be alleviated by inclusion of epitopes from multiple antigens from that pathogen in a vaccine composition. A “pathogen” may be an infectious agent or a tumor associated molecule.

One of the most formidable obstacles to the development of broadly efficacious epitope-based immunotherapeutics, however, has been the extreme polymorphism of HLA molecules. To date, effective non-genetically biased coverage of a population has been a task of considerable complexity; such coverage has required that epitopes be used

that are specific for HLA molecules corresponding to each individual HLA allele, therefore, impractically large numbers of epitopes would have to be used in order to cover ethnically diverse populations. Thus, there has existed a need for peptide epitopes that are bound by multiple HLA antigen molecules for use in epitope-based vaccines. The 5 greater the number of HLA antigen molecules bound, the greater the breadth of population coverage by the vaccine.

Furthermore, as described herein in greater detail, a need has existed to modulate peptide binding properties, for example, so that peptides that are able to bind to multiple HLA antigens do so with an affinity that will stimulate an immune response.

10 Identification of epitopes restricted by more than one HLA allele at an affinity that correlates with immunogenicity is important to provide thorough population coverage, and to allow the elicitation of responses of sufficient vigor to prevent or clear an infection in a diverse segment of the population. Such a response can also target a broad array of epitopes. The technology disclosed herein provides for such favored immune responses.

15 In a preferred embodiment, epitopes for inclusion in vaccine compositions of the invention are selected by a process whereby protein sequences of known antigens are evaluated for the presence of motif or supermotif-bearing epitopes. Peptides corresponding to a motif- or supermotif-bearing epitope are then synthesized and tested for the ability to bind to the HLA molecule that recognizes the selected motif. Those 20 peptides that bind at an intermediate or high affinity *i.e.*, an IC_{50} (or a K_D value) of 500 nM or less for HLA class I molecules or 1000 nM or less for HLA class II molecules, are further evaluated for their ability to induce a CTL or HTL response. Immunogenic peptide epitopes are selected for inclusion in vaccine compositions.

25 Supermotif-bearing peptides may additionally be tested for the ability to bind to multiple alleles within the HLA supertype family. Moreover, peptide epitopes may be analogued to modify binding affinity and/or the ability to bind to multiple alleles within an HLA supertype.

The invention also includes an embodiment comprising a method for monitoring or evaluating an immune response to HCV in a patient having a known HLA-type, the 30 method comprising incubating a T lymphocyte sample from the patient with a peptide composition comprising an HCV epitope consisting essentially of an amino acid sequence described in Tables VII to Table XX or Table XXII which binds the product of at least one HLA allele present in said patient, and detecting for the presence of a T lymphocyte

that binds to the peptide. A CTL peptide epitope may, for example, comprise a tetrameric complex.

An alternative modality for defining the peptide epitopes in accordance with the invention is to recite the physical properties, such as length; primary structure; or charge,

5 which are correlated with binding to a particular allele-specific HLA molecule or group of allele-specific HLA molecules. A further modality for defining peptide epitopes is to recite the physical properties of an HLA binding pocket, or properties shared by several allele-specific HLA binding pockets (*e.g.* pocket configuration and charge distribution) and reciting that the peptide epitope fits and binds to said pocket or pockets.

10 As will be apparent from the discussion below, other methods and embodiments are also contemplated. Further, novel synthetic peptides produced by any of the methods described herein are also part of the invention.

III. BRIEF DESCRIPTION OF THE FIGURES

15 Figure 1: Figure 1 provides a graph of total frequency of genotypes as a function of the number of HCV candidate epitopes bound by HLA-A and B molecules, in an average population.

Figure 2: Figure 2 illustrates the position of peptide epitopes in an experimental model minigene construct.

20

IV. DETAILED DESCRIPTION OF THE INVENTION

The peptide epitopes and corresponding nucleic acid compositions of the present invention are useful for stimulating an immune response to HCV by stimulating the production of CTL or HTL responses. The peptide epitopes, which are derived directly or 25 indirectly from native HCV amino acid sequences, are able to bind to HLA molecules and stimulate an immune response to HCV. The complete polyprotein sequence from HCV and its variants can be obtained from Genbank. Peptide epitopes and analogs thereof can also be readily determined from sequence information that may subsequently be discovered for heretofore unknown variants of HCV, as will be clear from the disclosure 30 provided below.

The peptide epitopes of the invention have been identified in a number of ways, as will be discussed below. Also discussed in greater detail is that analog peptides have been derived and the binding activity for HLA molecules modulated by modifying specific amino acid residues to create peptide analogs exhibiting altered immunogenicity.

Further, the present invention provides compositions and combinations of compositions that enable epitope-based vaccines that are capable of interacting with HLA molecules encoded by various genetic alleles to provide broader population coverage than prior vaccines.

5

IV.A. Definitions

The invention can be better understood with reference to the following definitions, which are listed alphabetically:

10 A "computer" or "computer system" generally includes: a processor; at least one information storage/retrieval apparatus such as, for example, a hard drive, a disk drive or a tape drive; at least one input apparatus such as, for example, a keyboard, a mouse, a touch screen, or a microphone; and display structure. Additionally, the computer may include a communication channel in communication with a network. Such a computer may include more or less than what is listed above.

15 "Cross-reactive binding" indicates that a peptide is bound by more than one HLA molecule; a synonym is degenerate binding.

A "cryptic epitope" elicits a response by immunization with an isolated peptide, but the response is not cross-reactive *in vitro* when intact whole protein which comprises the epitope is used as an antigen.

20 A "dominant epitope" is an epitope that induces an immune response upon immunization with a whole native antigen (see, e.g., Sercarz, *et al.*, *Annu. Rev. Immunol.* 11:729-766, 1993). Such a response is cross-reactive *in vitro* with an isolated peptide epitope.

25 With regard to a particular amino acid sequence, an "epitope" is a set of amino acid residues which is involved in recognition by a particular immunoglobulin, or in the context of T cells, those residues necessary for recognition by T cell receptor proteins and/or Major Histocompatibility Complex (MHC) receptors. In an immune system setting, *in vivo* or *in vitro*, an epitope is the collective features of a molecule, such as primary, secondary and tertiary peptide structure, and charge, that together form a site 30 recognized by an immunoglobulin, T cell receptor or HLA molecule. Throughout this disclosure epitope and peptide are often used interchangeably.

It is to be appreciated that protein or peptide molecules that comprise an epitope of the invention as well as additional amino acid(s) are still within the bounds of the invention. In certain embodiments, there is a limitation on the length of a peptide of the

invention which is not otherwise a construct. An embodiment that is length-limited occurs when the protein/peptide comprising an epitope of the invention comprises a region (i.e., a contiguous series of amino acids) having 100% identity with a native sequence. In order to avoid the definition of epitope from reading, e.g., on whole natural molecules, there is a limitation on the length of any region that has 100% identity with a native peptide sequence. Thus, for a peptide comprising an epitope of the invention and a region with 100% identity with a native peptide sequence (and is not otherwise a construct), the region with 100% identity to a native sequence generally has a length of: less than or equal to 600 amino acids, often less than or equal to 500 amino acids, often less than or equal to 400 amino acids, often less than or equal to 250 amino acids, often less than or equal to 100 amino acids, often less than or equal to 85 amino acids, often less than or equal to 75 amino acids, often less than or equal to 65 amino acids, and often less than or equal to 50 amino acids. In certain embodiments, an "epitope" of the invention is comprised by a peptide having a region with less than 51 amino acids that has 100% identity to a native peptide sequence, in any increment of (49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5) down to 5 amino acids.

Accordingly, peptide or protein sequences longer than 600 amino acids are within the scope of the invention, so long as they do not comprise any contiguous sequence of more than 600 amino acids that have 100% identity with a native peptide sequence, if they are not otherwise a construct. For any peptide that has five contiguous residues or less that correspond to a native sequence, there is no limitation on the maximal length of that peptide in order to fall within the scope of the invention. It is presently preferred that a CTL epitope be less than 600 residues long in any increment down to eight amino acid residues.

"Human Leukocyte Antigen" or "HLA" is a human class I or class II Major Histocompatibility Complex (MHC) protein (see, e.g., Stites, *et al.*, IMMUNOLOGY, 8TH ED., Lange Publishing, Los Altos, CA (1994).

An "HLA supertype or family", as used herein, describes sets of HLA molecules grouped on the basis of shared peptide-binding specificities. HLA class I molecules that share somewhat similar binding affinity for peptides bearing certain amino acid motifs are grouped into HLA supertypes. The terms HLA superfamily, HLA supertype family, HLA family, and HLA xx-like supertype molecules (where xx denotes a particular HLA type), are synonyms.

Throughout this disclosure, results are expressed in terms of "IC₅₀'s." IC₅₀ is the concentration of peptide in a binding assay at which 50% inhibition of binding of a reference peptide is observed. Given the conditions in which the assays are run (*i.e.*, limiting HLA proteins and labeled peptide concentrations), these values approximate K_D values. Assays for determining binding are described in detail, *e.g.*, in PCT publications WO 94/20127 and WO 94/03205. It should be noted that IC₅₀ values can change, often dramatically, if the assay conditions are varied, and depending on the particular reagents used (*e.g.*, HLA preparation, *etc.*). For example, excessive concentrations of HLA molecules will increase the apparent measured IC₅₀ of a given ligand.

10 Alternatively, binding is expressed relative to a reference peptide. Although as a particular assay becomes more, or less, sensitive, the IC₅₀'s of the peptides tested may change somewhat, the binding relative to the reference peptide will not significantly change. For example, in an assay run under conditions such that the IC₅₀ of the reference peptide increases 10-fold, the IC₅₀ values of the test peptides will also shift approximately 15 10-fold. Therefore, to avoid ambiguities, the assessment of whether a peptide is a good, intermediate, weak, or negative binder is generally based on its IC₅₀, relative to the IC₅₀ of a standard peptide.

Binding may also be determined using other assay systems including those using: live cells (*e.g.*, Ceppellini *et al.*, *Nature* 339:392, 1989; Christnick *et al.*, *Nature* 352:67, 20 1991; Busch *et al.*, *Int. Immunol.* 2:443, 1990; Hill *et al.*, *J. Immunol.* 147:189, 1991; del Guercio *et al.*, *J. Immunol.* 154:685, 1995), cell free systems using detergent lysates (*e.g.*, Cerundolo *et al.*, *J. Immunol.* 21:2069, 1991), immobilized purified MHC (*e.g.*, Hill *et al.*, *J. Immunol.* 152, 2890, 1994; Marshall *et al.*, *J. Immunol.* 152:4946, 1994), ELISA systems (*e.g.*, Reay *et al.*, *EMBO J.* 11:2829, 1992), surface plasmon resonance (*e.g.*, 25 Khilko *et al.*, *J. Biol. Chem.* 268:15425, 1993); high flux soluble phase assays (Hammer *et al.*, *J. Exp. Med.* 180:2353, 1994), and measurement of class I MHC stabilization or assembly (*e.g.*, Ljunggren *et al.*, *Nature* 346:476, 1990; Schumacher *et al.*, *Cell* 62:563, 1990; Townsend *et al.*, *Cell* 62:285, 1990; Parker *et al.*, *J. Immunol.* 149:1896, 1992).

As used herein, "high affinity" with respect to HLA class I molecules is defined as 30 binding with an IC₅₀, or K_D value, of 50 nM or less; "intermediate affinity" is binding with an IC₅₀ or K_D value of between about 50 and about 500 nM. "High affinity" with respect to binding to HLA class II molecules is defined as binding with an IC₅₀ or K_D value of 100 nM or less; "intermediate affinity" is binding with an IC₅₀ or K_D value of between about 100 and about 1000 nM.

The terms "identical" or percent "identity," in the context of two or more peptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using a sequence comparison algorithm or by manual alignment and visual inspection.

An "immunogenic peptide" or "peptide epitope" is a peptide that comprises an allele-specific motif or supermotif such that the peptide will bind an HLA molecule and induce a CTL and/or HTL response. Thus, immunogenic peptides of the invention are capable of binding to an appropriate HLA molecule and thereafter inducing an HLA-restricted cytotoxic or helper T cell response to the antigen from which the immunogenic peptide is derived.

The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany the material as it is found in its native state. Thus, isolated peptides in accordance with the invention preferably do not contain materials normally associated with the peptides in their *in situ* environment. An "isolated" epitope refers to an epitope that does not include the whole sequence of the antigen or polypeptide from which the epitope was derived. Typically the "isolated" epitope does not have attached thereto additional amino acids that result in a sequence that has 100% identity with a native sequence. The native sequence can be a sequence such as a tumor-associated antigen from which the epitope is derived.

"Major Histocompatibility Complex" or "MHC" is a cluster of genes that plays a role in control of the cellular interactions responsible for physiologic immune responses. In humans, the MHC complex is also known as the HLA complex. For a detailed description of the MHC and HLA complexes, see, Paul, FUNDAMENTAL IMMUNOLOGY, 25 3RD ED., Raven Press, New York, 1993.

The term "motif" refers to the pattern of residues in a peptide of defined length, usually a peptide of from about 8 to about 13 amino acids for a class I HLA motif and from about 6 to about 25 amino acids for a class II HLA motif, which is recognized by a particular HLA molecule. Peptide motifs are typically different for each protein encoded by each human HLA allele and differ in the pattern of the primary and secondary anchor residues.

A "negative binding residue" is an amino acid which, if present at certain positions (typically not primary anchor positions) in a peptide epitope, results in decreased binding affinity of the peptide for the peptide's corresponding HLA molecule.

A "non-native" sequence or "construct" refers to a sequence that is not found in nature ("non-naturally occurring"). Such sequences include, e.g., peptides that are lipidated or otherwise modified and polyepitopic compositions that contain epitopes that are non contiguous in a native protein sequence.

5 The term "peptide" is used interchangeably with "oligopeptide" in the present specification to designate a series of residues, typically L-amino acids, connected one to the other, typically by peptide bonds between the α -amino and carboxyl groups of adjacent amino acids. The preferred CTL-inducing peptides of the invention are 13 residues or less in length and usually consist of between about 8 and about 11 residues, 10 preferably 9 or 10 residues. The preferred HTL-inducing oligopeptides are less than about 50 residues in length and usually consist of between about 6 and about 30 residues, more usually between about 12 and 25, and often between about 15 and 20 residues.

"Pharmaceutically acceptable" refers to a generally non-toxic, inert, and/or physiologically compatible composition.

15 A "pharmaceutical excipient" comprises a material such as an adjuvant, a carrier, pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservative, and the like.

20 A "primary anchor residue" is an amino acid at a specific position along a peptide sequence which is understood to provide a contact point between the immunogenic peptide and the HLA molecule. One to three, usually two, primary anchor residues within a peptide of defined length generally defines a "motif" for an immunogenic peptide. These residues are understood to fit in close contact with peptide binding grooves of an HLA molecule, with their side chains buried in specific pockets of the binding grooves themselves. In one embodiment, the primary anchor residues are located 25 at position 2 (from the amino terminal position) and at the carboxyl terminal position of a 9-residue peptide epitope in accordance with the invention. The primary anchor positions for each motif and supermotif are set forth in Table 1. For example, analog peptides can be created by altering the presence or absence of particular residues in these primary anchor positions. Such analogs are used to modulate the binding affinity of a peptide 30 comprising a particular motif or supermotif.

"Promiscuous recognition" is where a distinct peptide is recognized by the same T cell clone in the context of various HLA molecules. Promiscuous recognition or binding is synonymous with cross-reactive binding.

A "protective immune response" or "therapeutic immune response" refers to a CTL and/or an HTL response to an antigen derived from an infectious agent or a tumor antigen, which prevents or at least partially arrests disease symptoms or progression. The immune response may also include an antibody response which has been facilitated by 5 the stimulation of helper T cells.

The term "residue" refers to an amino acid or amino acid mimetic incorporated into an oligopeptide by an amide bond or amide bond mimetic.

A "secondary anchor residue" is an amino acid at a position other than a primary anchor position in a peptide which may influence peptide binding. A secondary anchor 10 residue occurs at a significantly higher frequency amongst bound peptides than would be expected by random distribution of amino acids at one position. The secondary anchor residues are said to occur at "secondary anchor positions." A secondary anchor residue can be identified as a residue which is present at a higher frequency among high or intermediate affinity binding peptides, or a residue otherwise associated with high or 15 intermediate affinity binding. For example, analog peptides can be created by altering the presence or absence of particular residues in these secondary anchor positions. Such analogs are used to finely modulate the binding affinity of a peptide comprising a particular motif or supermotif.

A "subdominant epitope" is an epitope which evokes little or no response upon 20 immunization with whole antigens which comprise the epitope, but for which a response can be obtained by immunization with an isolated peptide, and this response (unlike the case of cryptic epitopes) is detected when whole protein is used to recall the response *in vitro* or *in vivo*.

A "supermotif" is a peptide binding specificity shared by HLA molecules encoded 25 by two or more HLA alleles. Preferably, a supermotif-bearing peptide is recognized with high or intermediate affinity (as defined herein) by two or more HLA antigens.

"Synthetic peptide" refers to a peptide that is man-made using such methods as chemical synthesis or recombinant DNA technology.

As used herein, a "vaccine" is a composition that contains one or more peptides of 30 the invention. There are numerous embodiments of vaccines in accordance with the invention, such as by a cocktail of one or more peptides; one or more epitopes of the invention comprised by a polyepitopic peptide; or nucleic acids that encode such peptides or polypeptides, e.g., a minigene that encodes a polyepitopic peptide. The "one or more peptides" can include, e.g., at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18,

19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50,
55, 60, 65, 70, 75, 80, 85, 90, 95 or 100 or more peptides of the invention. The peptides
or polypeptides can optionally be modified, such as by lipidation, addition of targeting or
other sequences. HLA class I-binding peptides of the invention can be admixed with, or
5 linked to, HLA class II-binding peptides, to facilitate activation of both cytotoxic T
lymphocytes and helper T lymphocytes. Vaccines can also comprise peptide-pulsed
antigen presenting cells, e.g., dendritic cells.

The nomenclature used to describe peptide compounds follows the conventional
practice wherein the amino group is presented to the left (the N-terminus) and the
10 carboxyl group to the right (the C-terminus) of each amino acid residue. When amino
acid residue positions are referred to in a peptide epitope they are numbered in an amino
to carboxyl direction with position one being the position closest to the amino terminal
end of the epitope, or the peptide or protein of which it may be a part. In the formulae
representing selected specific embodiments of the present invention, the amino- and
15 carboxyl-terminal groups, although not specifically shown, are in the form they would
assume at physiologic pH values, unless otherwise specified. In the amino acid structure
formulae, each residue is generally represented by standard three letter or single letter
designations. The L-form of an amino acid residue is represented by a capital single letter
or a capital first letter of a three-letter symbol, and the D-form for those amino acids
20 having D-forms is represented by a lower case single letter or a lower case three letter
symbol. Glycine has no asymmetric carbon atom and is simply referred to as "Gly" or G.
Symbols for the amino acids are shown below.

Single Letter Symbol	Three Letter Symbol	Amino Acids
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic Acid
E	Glu	Glutamic Acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

IV.B. Stimulation of CTL and HTL responses

The mechanism by which T cells recognize antigens has been delineated during the past ten years. Based on our understanding of the immune system we have developed efficacious peptide epitope vaccine compositions that can induce a therapeutic or prophylactic immune response to HCV in a broad population. For an understanding of the value and efficacy of the claimed compositions, a brief review of immunology-related technology is provided.

A complex of an HLA molecule and a peptidic antigen acts as the ligand recognized by HLA-restricted T cells (Buus, S. *et al.*, *Cell* 47:1071, 1986; Babbitt, B. P. *et al.*, *Nature* 317:359, 1985; Townsend, A. and Bodmer, H., *Annu. Rev. Immunol.* 7:601,

1989; Germain, R. N., *Annu. Rev. Immunol.* 11:403, 1993). Through the study of single amino acid substituted antigen analogs and the sequencing of endogenously bound, naturally processed peptides, critical residues that correspond to motifs required for specific binding to HLA antigen molecules have been identified and are described herein

5 and are set forth in Tables I, II, and III (see also, e.g., Southwood, *et al.*, *J. Immunol.* 160:3363, 1998; Rammensee, *et al.*, *Immunogenetics* 41:178, 1995; Rammensee *et al.*, SYFPEITHI, access via web at : <http://134.2.96.221/scripts.hlaserver.dll/home.htm>; Sette, A. and Sidney, J. *Curr. Opin. Immunol.* 10:478, 1998; Engelhard, V. H., *Curr. Opin. Immunol.* 6:13, 1994; Sette, A. and Grey, H. M., *Curr. Opin. Immunol.* 4:79, 1992;

10 Sinigaglia, F. and Hammer, J. *Curr. Biol.* 6:52, 1994; Ruppert *et al.*, *Cell* 74:929-937, 1993; Kondo *et al.*, *J. Immunol.* 155:4307-4312, 1995; Sidney *et al.*, *J. Immunol.* 157:3480-3490, 1996; Sidney *et al.*, *Human Immunol.* 45:79-93, 1996; Sette, A. and Sidney, J. *Immunogenetics*, in press, 1999).

Furthermore, x-ray crystallographic analysis of HLA-peptide complexes has

15 revealed pockets within the peptide binding cleft of HLA molecules which accommodate, in an allele-specific mode, residues borne by peptide ligands; these residues in turn determine the HLA binding capacity of the peptides in which they are present. (See, e.g., Madden, D.R. *Annu. Rev. Immunol.* 13:587, 1995; Smith, *et al.*, *Immunity* 4:203, 1996; Fremont *et al.*, *Immunity* 8:305, 1998; Stern *et al.*, *Structure* 2:245, 1994; Jones, E.Y.

20 *Curr. Opin. Immunol.* 9:75, 1997; Brown, J. H. *et al.*, *Nature* 364:33, 1993; Guo, H. C. *et al.*, *Proc. Natl. Acad. Sci. USA* 90:8053, 1993; Guo, H. C. *et al.*, *Nature* 360:364, 1992; Silver, M. L. *et al.*, *Nature* 360:367, 1992; Matsumura, M. *et al.*, *Science* 257:927, 1992; Madden *et al.*, *Cell* 70:1035, 1992; Fremont, D. H. *et al.*, *Science* 257:919, 1992; Saper, M. A. , Bjorkman, P. J. and Wiley, D. C., *J. Mol. Biol.* 219:277, 1991.)

25 Accordingly, the definition of class I and class II allele-specific HLA binding motifs, or class I or class II supermotifs allows identification of regions within a protein that have the potential of binding particular HLA antigen(s).

The present inventors have found that the correlation of binding affinity with immunogenicity, which is disclosed herein, is an important factor to be considered when

30 evaluating candidate peptides. Thus, by a combination of motif searches and HLA-peptide binding assays, candidates for epitope-based vaccines have been identified. After determining their binding affinity, additional confirmatory work can be performed to select, amongst these vaccine candidates, epitopes with preferred characteristics in terms of population coverage, antigenicity, and immunogenicity.

Various strategies can be utilized to evaluate immunogenicity, including:

- 1) Evaluation of primary T cell cultures from normal individuals (see, e.g., Wentworth, P. A. *et al.*, *Mol. Immunol.* 32:603, 1995; Celis, E. *et al.*, *Proc. Natl. Acad. Sci. USA* 91:2105, 1994; Tsai, V. *et al.*, *J. Immunol.* 158:1796, 1997; Kawashima, I. *et al.*, *Human Immunol.* 59:1, 1998); This procedure involves the stimulation of peripheral blood lymphocytes (PBL) from normal subjects with a test peptide in the presence of antigen presenting cells *in vitro* over a period of several weeks. T cells specific for the peptide become activated during this time and are detected using, e.g., a ^{51}Cr -release assay involving peptide sensitized target cells.
- 10 2) Immunization of HLA transgenic mice (see, e.g., Wentworth, P. A. *et al.*, *J. Immunol.* 26:97, 1996; Wentworth, P. A. *et al.*, *Int. Immunol.* 8:651, 1996; Alexander, J. *et al.*, *J. Immunol.* 159:4753, 1997); In this method, peptides in incomplete Freund's adjuvant are administered subcutaneously to HLA transgenic mice. Several weeks following immunization, splenocytes are removed and cultured *in vitro* in the presence of 15 test peptide for approximately one week. Peptide-specific T cells are detected using, e.g., a ^{51}Cr -release assay involving peptide sensitized target cells and target cells expressing endogenously generated antigen.
- 15 3) Demonstration of recall T cell responses from immune individuals who have effectively been vaccinated, recovered from infection, and/or from chronically infected 20 patients (see, e.g., Rehermann, B. *et al.*, *J. Exp. Med.* 181:1047, 1995; Doolan, D. L. *et al.*, *Immunity* 7:97, 1997; Bertoni, R. *et al.*, *J. Clin. Invest.* 100:503, 1997; Threlkeld, S. C. *et al.*, *J. Immunol.* 159:1648, 1997; Diepolder, H. M. *et al.*, *J. Virol.* 71:6011, 1997). In applying this strategy, recall responses are detected by culturing PBL from subjects 25 that have been naturally exposed to the antigen, for instance through infection, and thus have generated an immune response "naturally", or from patients who were vaccinated against the infection. PBL from subjects are cultured *in vitro* for 1-2 weeks in the presence of test peptide plus antigen presenting cells (APC) to allow activation of "memory" T cells, as compared to "naive" T cells. At the end of the culture period, T cell activity is detected using assays for T cell activity including ^{51}Cr release involving 30 peptide-sensitized targets, T cell proliferation, or lymphokine release.

The following describes the peptide epitopes and corresponding nucleic acids of the invention.

IV.C. Binding Affinity of Peptide Epitopes for HLA Molecules

The large degree of HLA polymorphism is an important factor to consider with the epitope-based approach to vaccine development. To address this factor, epitope selection including identification of peptides capable of binding at high or intermediate

5 affinity to multiple HLA molecules is often utilized, most preferably these epitopes bind at high or intermediate affinity to two or more allele specific HLA molecules.

CTL-inducing peptides of interest for vaccine compositions preferably include those that have an IC_{50} or binding affinity value for class I HLA molecules of 500 nM or better (*i.e.*, the value is \leq 500 nM). HTL-inducing peptides preferably include those that 10 have an IC_{50} or binding affinity value for class II HLA molecules of 1000 nM or better, (*i.e.*, the value is \leq 1,000 nM). For example, peptide binding is assessed by testing the capacity of a candidate peptide to bind to a purified HLA molecule *in vitro*. Peptides exhibiting high or intermediate affinity are then considered for further analysis. Selected peptides are tested on other members of the supertype family. In preferred embodiments, 15 peptides that exhibit cross-reactive binding are then used in vaccines or in cellular screening analyses.

Higher HLA binding affinity is typically correlated with greater immunogenicity. Greater immunogenicity can be manifested in several different ways. Immunogenicity corresponds to whether an immune response is elicited at all, and to the vigor of any 20 particular response, as well as to the extent of a population in which a response is elicited. For example, a peptide might elicit an immune response in a diverse array of the population, yet in no instance produce a vigorous response. In accordance with these principles, close to 90% of high binding peptides have been found to be immunogenic, as contrasted with about 50% of the peptides which bind with intermediate affinity. 25 Moreover, higher binding affinity peptides leads to more vigorous immunogenic responses. As a result, less peptide is required to elicit a similar biological effect if a high affinity binding peptide is used. Thus, in preferred embodiments of the invention, high affinity binding epitopes are particularly useful.

The relationship between binding affinity for HLA class I molecules and 30 immunogenicity of discrete peptide epitopes on bound antigens has been determined for the first time in the art by the present inventors. The correlation between binding affinity and immunogenicity was analyzed in two different experimental approaches (*see, e.g.*, Sette, *et al.*, *J. Immunol.* 153:5586-5592, 1994). In the first approach, the

immunogenicity of potential epitopes ranging in HLA binding affinity over a 10,000-fold range was analyzed in HLA-A*0201 transgenic mice. In the second approach, the antigenicity of approximately 100 different hepatitis B virus (HBV)-derived potential epitopes, all carrying A*0201 binding motifs, was assessed by using PBL from acute hepatitis patients. Pursuant to these approaches, it was determined that an affinity threshold value of approximately 500 nM (preferably 50 nM or less) determines the capacity of a peptide epitope to elicit a CTL response. These data are true for class I binding affinity measurements for naturally processed peptides and for synthesized T cell epitopes. These data also indicate the important role of determinant selection in the shaping of T cell responses (see, e.g., Schaeffer *et al. Proc. Natl. Acad. Sci. USA* 86:4649-4653, 1989).

An affinity threshold associated with immunogenicity in the context of HLA class II DR molecules has also been delineated (see, e.g., Southwood *et al. J. Immunology* 160:3363-3373, 1998). In order to define a biologically significant threshold of DR binding affinity, a database of the binding affinities of 32 DR-restricted epitopes for their restricting element (*i.e.*, the HLA molecule that binds the motif) was compiled. In approximately half of the cases (15 of 32 epitopes), DR restriction was associated with high binding affinities, *i.e.* binding affinity values of 100 nM or less. In the other half of the cases (16 of 32), DR restriction was associated with intermediate affinity (binding affinity values in the 100-1000 nM range). In only one of 32 cases was DR restriction associated with an IC_{50} of 1000 nM or greater. Thus, 1000 nM can be defined as an affinity threshold associated with immunogenicity in the context of DR molecules.

The binding affinity of peptides for HLA molecules can be determined as described in Example 1, below.

25

IV.D. Peptide Epitope Binding Motifs and Supermotifs

In the past few years evidence has accumulated to demonstrate that a large fraction of HLA class I and class II molecules can be classified into a relatively few supertypes, each characterized by largely overlapping peptide binding repertoires, and 30 consensus structures of the main peptide binding pockets.

For HLA molecule pocket analyses, the residues comprising the B and F pockets of HLA class I molecules as described in crystallographic studies were analyzed (see, e.g., Guo, H. C. *et al. Nature* 360:364, 1992; Saper, M. A. , Bjorkman, P. J. and Wiley, D. C., *J. Mol. Biol.* 219:277, 1991; Madden, D. R., Garboczi, D. N. and Wiley, D. C.,

Cell 75:693, 1993; Parham, P., Adams, E. J., and Arnett, K. L., *Immunol. Rev.* 143:141, 1995). In these analyses, residues 9, 45, 63, 66, 67, 70, and 99 were considered to make up the B pocket; and the B pocket was deemed to determine the specificity for the amino acid residue in the second position of peptide ligands. Similarly, residues 77, 80, 81, and 5 116 were considered to determine the specificity of the F pocket; the F pocket was deemed to determine the specificity for the C-terminal residue of a peptide ligand bound by the HLA class I molecule.

Through the study of single amino acid substituted antigen analogs and the sequencing of endogenously bound, naturally processed peptides, critical residues 10 required for allele-specific binding to HLA molecules have been identified. The presence of these residues correlates with binding affinity for HLA molecules. The identification of motifs and/or supermotifs that correlate with high and intermediate affinity binding is an important issue with respect to the identification of immunogenic peptide epitopes for the inclusion in a vaccine. Kast *et al.* (*J. Immunol.* 152:3904-3912, 1994) have shown 15 that motif-bearing peptides account for 90% of the epitopes that bind to allele-specific HLA class I molecules. In this study all possible peptides of 9 amino acids in length and overlapping by eight amino acids (240 peptides), which cover the entire sequence of the E6 and E7 proteins of human papillomavirus type 16, were evaluated for binding to five allele-specific HLA molecules that are expressed at high frequency among different 20 ethnic groups. This unbiased set of peptides allowed an evaluation of the predictive value of HLA class I motifs. From the set of 240 peptides, 22 peptides were identified that bound to an allele-specific HLA molecule with high or intermediate affinity. Of these 22 peptides, 20 (*i.e.* 91%) were motif-bearing. Thus, this study demonstrates the value of motifs for the identification of peptide epitopes for inclusion in a vaccine: application of 25 motif-based identification techniques eliminates screening of 90% of the potential epitopes in a target antigen protein sequence.

Such peptide epitopes are identified in the Tables described below.

Peptides of the present invention may also comprise epitopes that bind to MHC class II DR molecules. A greater degree of heterogeneity in both size and binding frame 30 position of the motif, relative to the N and C termini of the peptide, exists for class II peptide ligands. This increased heterogeneity of HLA class II peptide ligands is due to the structure of the binding groove of the HLA class II molecule which, unlike its class I counterpart, is open at both ends. Crystallographic analysis of HLA class II DRB*0101-peptide complexes showed that the major energy of binding is contributed by peptide

residues complexed with complementary pockets on the DRB*0101 molecules. An important anchor residue engages the deepest hydrophobic pocket (*see, e.g.*, Madden, D.R. *Ann. Rev. Immunol.* 13:587, 1995) and is referred to as position 1 (P1). P1 may represent the N-terminal residue of a class II binding peptide epitope, but more typically 5 is flanked towards the N-terminus by one or more residues. Other studies have also pointed to an important role for the peptide residue in the 6th position towards the C-terminus, relative to P1, for binding to various DR molecules.

Thus, peptides of the present invention are identified by any one of several HLA-specific amino acid motifs (*see, e.g.*, Tables I-III). If the presence of the motif 10 corresponds to the ability to bind several allele-specific HLA antigens, it is referred to as a supermotif. The HLA molecules that bind to peptides that possess a particular amino acid supermotif are collectively referred to as an HLA "supertype."

The peptide motifs and supermotifs described below, and summarized in Tables I-III, provide guidance for the identification and use of peptide epitopes in accordance with 15 the invention.

Examples of peptide epitopes bearing a respective supermotif or motif are included in Tables as designated in the description of each motif or supermotif below. The Tables include a binding affinity ratio listing for some of the peptide epitopes. The ratio may be converted to IC₅₀ by using the following formula: IC₅₀ of the standard 20 peptide/ratio = IC₅₀ of the test peptide (*i.e.*, the peptide epitope). The IC₅₀ values of standard peptides used to determine binding affinities for Class I peptides are shown in Table IV. The IC₅₀ values of standard peptides used to determine binding affinities for Class II peptides are shown in Table V. The peptides used as standards for the binding assays described herein are examples of standards; alternative standard peptides can also 25 be used when performing such an analysis.

To obtain the peptide epitope sequences listed in each Table, protein sequence data from fourteen HCV isolates were evaluated for the presence of the designated supermotif or motif. The fourteen strains include HPCCGAA, HPCPLYPRE, HCV-H-CMR, HCV-J1, HPCGENANTI, HPCGENOM, HPCHUMR, HPCJCG, HPCJTA, HCV-30 J483, HCV-JK1, HCV-N, HPCPOLP, and HCV-J8. Peptide epitopes were additionally evaluated on the basis of their conservancy among these fourteen strains. A criterion for conservancy requires that the entire sequence of an HLA class I binding peptide be totally conserved in 79% of the sequences available for a specific protein. Similarly, a criterion for conservancy requires that the entire 9-mer core region of an HLA class II binding

peptide be totally conserved in 79% of the sequences available for a specific protein. The percent conservancy of the selected peptide epitopes is indicated on the Tables. The frequency, *i.e.* the number of strains of the fourteen strains in which the totally conserved peptide sequence was identified, is also shown. The "position" column in the Tables 5 designates the amino acid position of the HCV polyprotein that corresponds to the first amino acid residue of the epitope. The "number of amino acids" indicates the number of residues in the epitope sequence.

HLA Class I Motifs Indicative of CTL Inducing Peptide Epitopes:

10 The primary anchor residues of the HLA class I peptide epitope supermotifs and motifs delineated below are summarized in Table I. The HLA class I motifs set out in Table I(a) are those most particularly relevant to the invention claimed here. Primary and secondary anchor positions are summarized in Table II. Allele-specific HLA molecules that comprise HLA class I supertype families are listed in Table VI.

15

IV.D.1. HLA-A1 supermotif

The HLA-A1 supermotif is characterized by the presence in peptide ligands of a small (T or S) or hydrophobic (L, I, V, or M) primary anchor residue in position 2, and an aromatic (Y, F, or W) primary anchor residue at the C-terminal position of the epitope.

20 The corresponding family of HLA molecules that bind to the A1 supermotif (*i.e.*, the HLA-A1 supertype) includes at least A*0101, A*2601, A*2602, A*2501, and A*3201 (see, e.g., DiBrino, M. *et al.*, *J. Immunol.* 151:5930, 1993; DiBrino, M. *et al.*, *J. Immunol.* 152:620, 1994; Kondo, A. *et al.*, *Immunogenetics* 45:249, 1997). Other allele-specific HLA molecules predicted to be members of the A1 superfamily are shown in 25 Table VI. Peptides binding to each of the individual HLA proteins can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the A1 supermotif are set forth in Table VII.

30 **IV.D.2. HLA-A2 supermotif**

Primary anchor specificities for allele-specific HLA-A2.1 molecules (Falk *et al.*, *Nature* 351:290-296, 1991; Hunt *et al.*, *Science* 255:1261-1263, 1992; Parker *et al.*, *J. Immunol.* 149:3580-3587, 1992) and cross-reactive binding within the HLA A2 family (Fruci *et al.*, *Human Immunol.* 38:187-192, 1993; Tanigaki *et al.*, *Human Immunol.*

39:155-162, 1994) have been described. The present inventors have defined additional primary anchor residues that determine cross-reactive binding to multiple allele-specific HLA A2 molecules (Ruppert *et al.*, *Cell* 74:929-937, 1993; Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994). The HLA-
5 A2 supermotif comprises peptide ligands with L, I, V, M, A, T, or Q as a primary anchor residue at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope.

The corresponding family of HLA molecules (*i.e.*, the HLA-A2 supertype that binds these peptides) is comprised of at least: A*0201, A*0202, A*0203, A*0204,
10 A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, and A*6901. Other allele-specific HLA molecules predicted to be members of the A2 superfamily are shown in Table VI. As explained in detail below, binding to each of the individual allele-specific HLA molecules can be modulated by substitutions at the primary anchor and/or secondary anchor positions, preferably choosing respective residues specified for the
15 supermotif.

Peptide epitopes that comprise an A2 supermotif are set forth in Table VIII. The motifs comprising the primary anchor residues V, A, T, or Q at position 2 and L, I, V, A, or T at the C-terminal position are those most particularly relevant to the invention claimed herein.

20

IV.D.3. HLA-A3 supermotif

The HLA-A3 supermotif is characterized by the presence in peptide ligands of A, L, I, V, M, S, or, T as a primary anchor at position 2, and a positively charged residue, R or K, at the C-terminal position of the epitope (*e.g.*, in position 9 of 9-mers). Exemplary
25 members of the corresponding family of HLA molecules (the HLA-A3 supertype) that bind the A3 supermotif include at least A*0301, A*1101, A*3101, A*3301, and A*6801. Other allele-specific HLA molecules predicted to be members of the A3 superfamily are shown in Table VI. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions of amino acids
30 at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the A3 supermotif are set forth in Table IX.

IV.D.4. HLA-A24 supermotif

The HLA-A24 supermotif is characterized by the presence in peptide ligands of an aromatic (F, W, or Y) or hydrophobic aliphatic (L, I, V, M, or T) residue as a primary anchor in position 2, and Y, F, W, L, I, or M as primary anchor at the C-terminal position 5 of the epitope. The corresponding family of HLA molecules that bind to the A24 supermotif (*i.e.*, the A24 supertype) includes at least A*2402, A*3001, and A*2301. Other allele-specific HLA molecules predicted to be members of the A24 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably 10 choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the A24 supermotif are set forth in Table X.

IV.D.5. HLA-B7 supermotif

The HLA-B7 supermotif is characterized by peptides bearing proline in position 2 15 as a primary anchor, and a hydrophobic or aliphatic amino acid (L, I, V, M, A, F, W, or Y) as the primary anchor at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind the B7 supermotif (*i.e.*, the HLA-B7 supertype) is comprised of at least twenty six HLA-B proteins including: B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3504, B*3505, B*3506, B*3507, 20 B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, and B*7801 (*see, e.g.*, Sidney, *et al.*, *J. Immunol.* 154:247, 1995; Barber, *et al.*, *Curr. Biol.* 5:179, 1995; Hill, *et al.*, *Nature* 360:434, 1992; 25 Rammensee, *et al.*, *Immunogenetics* 41:178, 1995). Other allele-specific HLA molecules predicted to be members of the B7 superfamily are shown in Table VI. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be 30 modulated by substitutions at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the B7 supermotif are set forth in Table XI.

IV.D.6. HLA-B27 supermotif

The HLA-B27 supermotif is characterized by the presence in peptide ligands of a positively charged (R, H, or K) residue as a primary anchor at position 2, and a hydrophobic (F, Y, L, W, M, I, A, or V) residue as a primary anchor at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA

molecules that bind to the B27 supermotif (*i.e.*, the B27 supertype) include at least B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*3801, B*3901, B*3902, and B*7301. Other allele-specific HLA molecules predicted to be members of the B27 superfamily are shown in Table VI. Peptide binding to each of the 5 allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the B27 supermotif are set forth in Table XII.

10 **IV.D.7. HLA-B44 supermotif**

The HLA-B44 supermotif is characterized by the presence in peptide ligands of negatively charged (D or E) residues as a primary anchor in position 2, and hydrophobic residues (F, W, Y, L, I, M, V, or A) as a primary anchor at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA molecules that bind to 15 the B44 supermotif (*i.e.*, the B44 supertype) include at least: B*1801, B*1802, B*3701, B*4001, B*4002, B*4006, B*4402, B*4403, and B*4006. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions; preferably choosing respective residues specified for the supermotif.

20

IV.D.8. HLA-B58 supermotif

The HLA-B58 supermotif is characterized by the presence in peptide ligands of a small aliphatic residue (A, S, or T) as a primary anchor residue at position 2, and an aromatic or hydrophobic residue (F, W, Y, L, I, V, M, or A) as a primary anchor residue 25 at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA molecules that bind to the B58 supermotif (*i.e.*, the B58 supertype) include at least: B*1516, B*1517, B*5701, B*5702, and B*5801. Other allele-specific HLA molecules predicted to be members of the B58 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by 30 substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the B58 supermotif are set forth in Table XIII.

IV.D.9. HLA-B62 supermotif

The HLA-B62 supermotif is characterized by the presence in peptide ligands of the polar aliphatic residue Q or a hydrophobic aliphatic residue (L, V, M, I, or P) as a primary anchor in position 2, and a hydrophobic residue (F, W, Y, M, I, V, L, or A) as a 5 primary anchor at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA molecules that bind to the B62 supermotif (*i.e.*, the B62 supertype) include at least: B*1501, B*1502, B*1513, and B5201. Other allele-specific HLA molecules predicted to be members of the B62 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by 10 substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the B62 supermotif are set forth in Table XIV.

IV.D.10. HLA-A1 motif

15 The HLA-A1 motif is characterized by the presence in peptide ligands of T, S, or M as a primary anchor residue at position 2 and the presence of Y as a primary anchor residue at the C-terminal position of the epitope. An alternative allele-specific A1 motif is characterized by a primary anchor residue at position 3 rather than position 2. This motif is characterized by the presence of D, E, A, or S as a primary anchor residue in 20 position 3, and a Y as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA A1 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

25 Peptide epitopes that comprise either A1 motif are set forth in Table XV. The epitopes comprising T, S, or M at position 2 and Y at the C-terminal position are also included in the listing of HLA-A1 supermotif-bearing peptide epitopes listed in Table VII.

IV.D.11. HLA-A*0201 motif

30 An HLA-A2*0201 motif was first determined to be characterized by the presence in peptide ligands of L or M as a primary anchor residue in position 2, and L or V as a primary anchor residue at the C-terminal position of a 9-residue peptide (Falk *et al.*, *Nature* 351:290-296, 1991). The A*0201 motif was also determined to further comprise an I at position 2 and I or A at the C-terminal position of a nine amino acid peptide (Hunt

et al., *Science* 255:1261-1263, March 6, 1992; Parker *et al.*, *J. Immunol.* 149:3580-3587, 1992). Subsequently, the A*0201 allele-specific motif has been defined by the present inventors to additionally comprise V, A, T, or Q as a primary anchor residue at position 2, and M as a primary anchor residue at the C-terminal position of the epitope.

5 Additionally, the A*0201 allele-specific motif has been found to comprise a T at the C-terminal position (Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994). Thus, the HLA-A*0201 motif comprises peptide ligands with L, I, V, M, A, T, or Q as primary anchor residues at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope. The preferred and tolerated residues that characterize the

10 primary anchor positions of the HLA-A*0201 motif are identical to the residues describing the A2 supermotif. (For reviews of relevant data, see, e.g., Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Ruppert *et al.*, *Cell* 74:929-937, 1993; Sidney *et al.*, *Immunol. Today* 17:261-266, 1996; Sette and Sidney, *Curr. Opin. in Immunol.* 10:478-482, 1998). Secondary anchor residues that characterize the A*0201 motif have

15 additionally been defined as disclosed herein. These are disclosed in Table II. Peptide binding to HLA-A*0201 molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Peptide epitopes that comprise an A*0201 motif are set forth in Table VIII. The
20 A*0201 motifs comprising the primary anchor residues V, A, T, or Q at position 2 and L, I, V, A, or T at the C-terminal position are those most particularly relevant to the invention claimed herein.

IV.D.12. HLA-A3 motif

25 The HLA-A3 motif is characterized by the presence in peptide ligands of L, M, V, I, S, A, T, F, C, G, or D as a primary anchor residue at position 2, and the presence of K, Y, R, H, F, or A as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA-A3 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the
30 motif.

The A3 supermotif primary anchor residues comprise a subset of the A3- and A11-allele specific motif primary anchor residues. Peptide epitopes that comprise the A3 motif are set forth in Table XVI. Those peptide epitopes that also comprise the A3 supermotif are also listed in Table IX.

IV.D.13. HLA-A11 motif

The HLA-A11 motif is characterized by the presence in peptide ligands of V, T, M, L, I, S, A, G, N, C, D, or F as a primary anchor residue in position 2, and K, R, Y, or 5 H as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA-A11 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Peptide epitopes that comprise the A11 motif are set forth in Table XVII; peptide epitopes comprising the A3 allele-specific motif are also present in this Table because of 10 the overlap between the A3 and A11 motif primary anchor specificities. Further, those peptide epitopes that comprise the A3 supermotif are also listed in Table IX.

IV.D.14. HLA-A24 motif

The HLA-A24 motif is characterized by the presence in peptide ligands of Y, F, 15 W, or M as a primary anchor residue in position 2, and F, L, I, or W as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA-A24 molecules can be modulated by substitutions at primary and/or secondary anchor positions; preferably choosing respective residues specified for the motif.

Peptide epitopes that comprise the A24 motif are set forth in Table XVIII. These 20 epitopes are also listed in Table X, which sets forth HLA-A24-supermotif-bearing peptide epitopes, as the primary anchor residues characterizing the A24 allele-specific motif comprise a subset of the A24 supermotif primary anchor residues.

HLA Class II Binding Motifs

25 The primary and secondary anchor residues of the HLA class II peptide epitope supermotifs and motifs delineated below are summarized in Table III.

IV.D.15. HLA DR-1-4-7 supermotif

Motifs have also been identified for peptides that bind to three common HLA 30 class II allele-specific HLA molecules: HLA DRB1*0401, DRB1*0101, and DRB1*0701. Collectively, the common residues from these motifs delineate the HLA DR-1-4-7 supermotif. Peptides that bind to these DR molecules carry a supermotif characterized by a large aromatic or hydrophobic residue (Y, F, W, L, I, V, or M) as a primary anchor residue in position 1, and a small, non-charged residue (S, T, C, A, P, V,

I, L, or M) as a primary anchor residue in position 6 of a 9-mer core region. Allele-specific secondary effects and secondary anchors for each of these HLA types have also been identified. These are set forth in Table III. Peptide binding to HLA- DRB1*0401, DRB1*0101, and/or DRB1*0701 can be modulated by substitutions at primary and/or 5 secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Conserved peptide epitopes *i.e.*, conserved in $\geq 79\%$ ($\geq 11/14$) of the HCV strains used for the present analysis, may be described as corresponding to epitopes containing a nine residue core comprising the DR-1-4-7 supermotif, and in which the 9 residue core is 10 conserved in $\geq 79\%$ (wherein position 1 of the motif is at position 1 of the nine residue core). Conserved 9-mer core regions are set forth in Table XIXa. Respective exemplary peptide epitopes of 15 amino acid residues in length, each of which comprise a conserved nine residue core, are also shown in section "a" of the Table. Cross-reactive binding data for exemplary 15-residue supermotif-bearing peptides are shown in Table XIXb.

15

IV.D.16. HLA DR3 motifs

Two alternative motifs (*i.e.*, submotifs) characterize peptide epitopes that bind to HLA-DR3 molecules. In the first motif (submotif DR3A) a large, hydrophobic residue (L, I, V, M, F, or Y) is present in anchor position 1 of a 9-mer core, and D is present as an 20 anchor at position 4, towards the carboxyl terminus of the epitope. As in other class II motifs, core position 1 may or may not occupy the peptide N-terminal position.

The alternative DR3 submotif provides for lack of the large, hydrophobic residue at anchor position 1, and/or lack of the negatively charged or amide-like anchor residue at position 4, by the presence of a positive charge at position 6 towards the carboxyl 25 terminus of the epitope. Thus, for the alternative allele-specific DR3 motif (submotif DR3B): L, I, V, M, F, Y, A, or Y is present at anchor position 1; D, N, Q, E, S, or T is present at anchor position 4; and K, R, or H is present at anchor position 6. Peptide binding to HLA-DR3 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

30

Conserved 9-mer core regions (*i.e.*, those sequences that are conserved in at least 79% of the 14 HCV strains used for the analysis) corresponding to a nine residue sequence comprising the DR3A submotif (wherein position 1 of the motif is at position 1 of the nine residue core) are set forth in Table XXa. Respective exemplary peptide

epitopes of 15 amino acid residues in length, each of which comprise a conserved nine residue core, are also shown in Table XXa. Table XXb shows binding data of exemplary DR3 submotif A-bearing peptides.

Conserved 9-mer core regions (*i.e.*, those that are at least 79% conserved in the 14 5 HCV strains used for the analysis) comprising the DR3B submotif and respective exemplary 15-mer peptides comprising the DR3 submotif-B epitope are set forth in Table XXc. Table XXd shows binding data of exemplary DR3 submotif B-bearing peptides.

Each of the HLA class I or class II peptide epitopes set out in the Tables herein 10 are deemed singly to be an inventive aspect of this application. Further, it is also an inventive aspect of this application that each peptide epitope may be used in combination with any other peptide epitope.

IV.E. Enhancing Population Coverage of the Vaccine

Vaccines that have broad population coverage are preferred because they are more 15 commercially viable and generally applicable to the most people. Broad population coverage can be obtained using the peptides of the invention (and nucleic acid compositions that encode such peptides) through selecting peptide epitopes that bind to HLA alleles which, when considered in total, are present in most of the population. Table XXI lists the overall frequencies of the HLA class I supertypes in various ethnicities 20 (Table XXIa) and the combined population coverage achieved by the A2-, A3-, and B7- supertypes (Table XXIb). The A2-, A3-, and B7 supertypes are each present on the average of over 40% in each of these five major ethnic groups. Coverage in excess of 80% is achieved with a combination of these supermotifs. These results suggest that 25 effective and non-ethnically biased population coverage is achieved upon use of a limited number of cross-reactive peptides. Although the population coverage reached with these three main peptide specificities is high, coverage can be expanded to reach 95% population coverage and above, and more easily achieve truly multispecific responses upon use of additional supermotif or allele-specific motif bearing peptides.

The B44-, A1-, and A24-supertypes are present, on average, in a range from 25% 30 to 40% in these major ethnic populations (Table XXIa). While less prevalent overall, the B27-, B58-, and B62 supertypes are each present with a frequency >25% in at least one major ethnic group (Table XXIa). Table XXIb summarizes the estimated prevalence of combinations of HLA supertypes that have been identified in five major ethnic groups.

The incremental coverage obtained by the inclusion of A1,- A24-, and B44-supertypes to the A2, A3, and B7 coverage, or all of the supertypes described herein, is shown.

The data presented herein, together with the previous definition of the A2-, A3-, and B7-supertypes, indicates that all antigens, with the possible exception of A29, B8, 5 and B46, can be classified into a total of nine HLA supertypes. By including epitopes from the six most frequent supertypes, an average population coverage of 99% is obtained for five major ethnic groups..

IV.F. Immune Response-Stimulating Peptide Analogs

10 In general, CTL and HTL responses are not directed against all possible epitopes. Rather, they are restricted to a few "immunodominant" determinants (Zinkernagel, *et al.*, *Adv. Immunol.* 27:5159, 1979; Bennink, *et al.*, *J. Exp. Med.* 168:19351939, 1988; Rawle, *et al.*, *J. Immunol.* 146:3977-3984, 1991). It has been recognized that immunodominance (Benacerraf, *et al.*, *Science* 175:273-279, 1972) could be explained by either the ability of 15 a given epitope to selectively bind a particular HLA protein (determinant selection theory) (Vitiello, *et al.*, *J. Immunol.* 131:1635, 1983); Rosenthal, *et al.*, *Nature* 267:156-158, 1977), or to be selectively recognized by the existing TCR (T cell receptor) specificities (repertoire theory) (Klein, J., *IMMUNOLOGY, THE SCIENCE OF SELFNONSELF DISCRIMINATION*, John Wiley & Sons, New York, pp. 270-310, 1982). It has been 20 demonstrated that additional factors, mostly linked to processing events, can also play a key role in dictating, beyond strict immunogenicity, which of the many potential determinants will be presented as immunodominant (Sercarz, *et al.*, *Annu. Rev. Immunol.* 11:729-766, 1993).

The concept of dominance and subdominance is relevant to immunotherapy of 25 both infectious diseases and cancer. For example, in the course of chronic viral disease, recruitment of subdominant epitopes can be important for successful clearance of the infection, especially if dominant CTL or HTL specificities have been inactivated by functional tolerance, suppression, mutation of viruses and other mechanisms (Franco, *et al.*, *Curr. Opin. Immunol.* 7:524-531, 1995). In the case of cancer and tumor antigens, 30 CTLs recognizing at least some of the highest binding affinity peptides might be functionally inactivated. Lower binding affinity peptides are preferentially recognized at these times, and may therefore be preferred in therapeutic or prophylactic anti-cancer vaccines.

In particular, it has been noted that a significant number of epitopes derived from known non-viral tumor associated antigens (TAA) bind HLA class I with intermediate affinity (IC_{50} in the 50-500 nM range). For example, it has been found that 8 of 15 known TAA peptides recognized by tumor infiltrating lymphocytes (TIL) or CTL bound 5 in the 50-500 nM range. (These data are in contrast with estimates that 90% of known viral antigens were bound by HLA class I molecules with IC_{50} of 50 nM or less, while only approximately 10% bound in the 50-500 nM range (Sette, *et al.*, *J. Immunol.*, 153:558-5592, 1994). In the cancer setting this phenomenon is probably due to 10 elimination or functional inhibition of the CTL recognizing several of the highest binding peptides, presumably because of T cell tolerization events.

Without intending to be bound by theory, it is believed that because T cells to dominant epitopes may have been clonally deleted, selecting subdominant epitopes may allow existing T cells to be recruited, which will then lead to a therapeutic or prophylactic response. However, the binding of HLA molecules to subdominant epitopes is often less 15 vigorous than to dominant ones. Accordingly, there is a need to be able to modulate the binding affinity of particular immunogenic epitopes for one or more HLA molecules, and thereby to modulate the immune response elicited by the peptide, for example to prepare analog peptides which elicit a more vigorous response. This ability would greatly enhance the usefulness of peptide-based vaccines and therapeutic agents.

20 Although peptides with suitable cross-reactivity among all alleles of a superfamily are identified by the screening procedures described above, cross-reactivity is not always as complete as possible, and in certain cases procedures to increase cross-reactivity of peptides can be useful; moreover, such procedures can also be used to modify other properties of the peptides such as binding affinity or peptide stability. Having established 25 the general rules that govern cross-reactivity of peptides for HLA alleles within a given motif or supermotif, modification (*i.e.*, analoging) of the structure of peptides of particular interest in order to achieve broader (or otherwise modified) HLA binding capacity can be performed. More specifically, peptides which exhibit the broadest cross-reactivity patterns, can be produced in accordance with the teachings herein. The present 30 concepts related to analog generation are set forth in greater detail in co-pending U.S.S.N. 09/226,775 filed 1/6/99.

In brief, the strategy employed utilizes the motifs or supermotifs which correlate with binding to certain HLA molecules. The motifs or supermotifs are defined by having primary anchors, and in many cases secondary anchors. Analog peptides can be created

by substituting amino acid residues at primary anchor, secondary anchor, or at primary and secondary anchor positions. Generally, analogs are made for peptides that already bear a motif or supermotif. Preferred secondary anchor residues of supermotifs and motifs that have been defined for HLA class I and class II binding peptides are shown in 5 Tables II and III, respectively.

For a number of the motifs or supermotifs in accordance with the invention, residues are defined which are deleterious to binding to allele-specific HLA molecules or members of HLA supertypes that bind the respective motif or supermotif (Tables II and III). Accordingly, removal of such residues that are detrimental to binding can be 10 performed in accordance with the present invention. For example, in the case of the A3 supertype, when all peptides that have such deleterious residues are removed from the population of analyzed peptides, the incidence of cross-reactivity increases from 22% to 37% (see, e.g., Sidney, J. et al., *Hu. Immunol.* 45:79, 1996). Thus, one strategy to improve the cross-reactivity of peptides within a given supermotif is simply to delete one 15 or more of the deleterious residues present within a peptide and substitute a small “neutral” residue such as Ala (that may not influence T cell recognition of the peptide). An enhanced likelihood of cross-reactivity is expected if, together with elimination of detrimental residues within a peptide, “preferred” residues associated with high affinity binding to an allele-specific HLA molecule or to multiple HLA molecules within a 20 superfamily are inserted.

To ensure that an analog peptide, when used as a vaccine, actually elicits a CTL response to the native epitope *in vivo* (or, in the case of class II epitopes, elicits helper T cells that cross-react with the wild type peptides), the analog peptide may be used to immunize T cells *in vitro* from individuals of the appropriate HLA allele. Thereafter, the 25 immunized cells' capacity to induce lysis of wild type peptide sensitized target cells is evaluated. It will be desirable to use as antigen presenting cells, cells that have been either infected, or transfected with the appropriate genes, or, in the case of class II epitopes only, cells that have been pulsed with whole protein antigens, to establish whether endogenously produced antigen is also recognized by the relevant T cells.

30 Another embodiment of the invention is to create analogs of weak binding peptides, to thereby ensure adequate numbers of cross-reactive cellular binders. Class I binding peptides exhibiting binding affinities of 500-5000 nM, and carrying an acceptable but suboptimal primary anchor residue at one or both positions can be “fixed” by

substituting preferred anchor residues in accordance with the respective supertype. The analog peptides can then be tested for crossbinding activity.

Another embodiment for generating effective peptide analogs involves the substitution of residues that have an adverse impact on peptide stability or solubility in, 5 *e.g.*, a liquid environment. This substitution may occur at any position of the peptide epitope. For example, a cysteine (C) can be substituted out in favor of α -amino butyric acid. Due to its chemical nature, cysteine has the propensity to form disulfide bridges and sufficiently alter the peptide structurally so as to reduce binding capacity. Substituting α -amino butyric acid for C not only alleviates this problem, but actually improves binding 10 and crossbinding capability in certain instances (*see, e.g.*, the review by Sette *et al.*, In: Persistent Viral Infections, Eds. R. Ahmed and I. Chen, John Wiley & Sons, England, 1999). Substitution of cysteine with α -amino butyric acid may occur at any residue of a peptide epitope, *i.e.* at either anchor or non-anchor positions.

15 Representative analog peptides are set forth in Table XXII. The Table indicates the length and sequence of the analog peptide as well as the motif or supermotif, if appropriate. The information in the "Fixed Nomenclature" column indicates the residues substituted at the indicated position numbers for the respective analog.

IV.G. Computer Screening of Protein Sequences from Disease-Related Antigens for 20 Supermotif- or Motif-Bearing Peptides

In order to identify supermotif- or motif-bearing epitopes in a target antigen, a native protein sequence, *e.g.*, a tumor-associated antigen, or sequences from an infectious organism, or a donor tissue for transplantation, is screened using a means for computing, such as an intellectual calculation or a computer, to determine the presence of a 25 supermotif or motif within the sequence. The information obtained from the analysis of native peptide can be used directly to evaluate the status of the native peptide or may be utilized subsequently to generate the peptide epitope.

Computer programs that allow the rapid screening of protein sequences for the occurrence of the subject supermotifs or motifs are encompassed by the present 30 invention; as are programs that permit the generation of analog peptides. These programs are implemented to analyze any identified amino acid sequence or operate on an unknown sequence and simultaneously determine the sequence and identify motif-bearing epitopes thereof; analogs can be simultaneously determined as well. Generally, the identified sequences will be from a pathogenic organism or a tumor-associated peptide. For

example, the target molecules considered herein include, without limitation, the core, S, E1, NS1/E2, NS2, NS3, NS4, and NS5 regions of HCV.

In cases where the sequence of multiple variants of the same target protein are available, peptides may also be selected on the basis of their conservancy. A presently preferred criterion for conservancy defines that the entire sequence of an HLA class I binding peptide or the entire 9-mer core of a class II binding peptide, be totally (*i.e.*, 100%) conserved in at least 79% of the sequences evaluated for a specific protein. This definition of conservancy has been employed herein; although, as appreciated by those in the art, lower or higher degrees of conservancy can be employed as appropriate for a given antigenic target.

It is important that the selection criteria utilized for prediction of peptide binding are as accurate as possible, to correlate most efficiently with actual binding. Prediction of peptides that bind, for example, to HLA-A*0201, on the basis of the presence of the appropriate primary anchors, is positive at about a 30% rate (see, *e.g.*, Ruppert, J. *et al.* 15 *Cell* 74:929, 1993). However, by extensively analyzing peptide-HLA binding data disclosed herein, data in related patent applications, and data in the art, the present inventors have developed a number of allele-specific polynomial algorithms that dramatically increase the predictive value over identification on the basis of the presence of primary anchor residues alone. These algorithms take into account not only the 20 presence or absence of primary anchors, but also consider the positive or deleterious presence of secondary anchor residues (to account for the impact of different amino acids at different positions). The algorithms are essentially based on the premise that the overall affinity (or ΔG) of peptide-HLA interactions can be approximated as a linear polynomial function of the type:

$$25 \quad \Delta G = a_{1i} \times a_{2i} \times a_{3i} \dots \times a_{ni}$$

where a_{ji} is a coefficient that represents the effect of the presence of a given amino acid (j) at a given position (i) along the sequence of a peptide of n amino acids. An important assumption of this method is that the effects at each position are essentially independent of each other. This assumption is justified by studies that demonstrated that peptides are 30 bound to HLA molecules and recognized by T cells in essentially an extended conformation. Derivation of specific algorithm coefficients has been described, for example, in Gulukota, K. *et al.*, *J. Mol. Biol.* 267:1258, 1997.

Additional methods to identify preferred peptide sequences, which also make use of specific motifs, include the use of neural networks and molecular modeling programs (see, e.g., Milik *et al.*, *Nature Biotechnology* 16:753, 1998; Altuvia *et al.*, *Hum. Immunol.* 58:1, 1997; Altuvia *et al.*, *J. Mol. Biol.* 249:244, 1995; Buus, S. *Curr. Opin. Immunol.* 11:209-213, 1999; Brusic, V. *et al.*, *Bioinformatics* 14:121-130, 1998; Parker *et al.*, *J. Immunol.* 152:163, 1993; Meister *et al.*, *Vaccine* 13:581, 1995; Hammer *et al.*, *J. Exp. Med.* 180:2353, 1994; Sturniolo *et al.*, *Nature Biotechnol.* 17:555 1999).

For example, it has been shown that in sets of A*0201 motif-bearing peptides containing at least one preferred secondary anchor residue while avoiding the presence of any deleterious secondary anchor residues, 69% of the peptides will bind A*0201 with an IC₅₀ less than 500 nM (Ruppert, J. *et al.* *Cell* 74:929, 1993). These algorithms are also flexible in that cut-off scores may be adjusted to select sets of peptides with greater or lower predicted binding properties, as desired.

In utilizing computer screening to identify peptide epitopes, a protein sequence or translated sequence may be analyzed using software developed to search for motifs, for example the "FINDPATTERNS" program (Devereux, *et al.* *Nucl. Acids Res.* 12:387-395, 1984) or MotifSearch 1.4 software program (D. Brown, San Diego, CA) to identify potential peptide sequences containing appropriate HLA binding motifs. The identified peptides can be scored using customized polynomial algorithms to predict their capacity to bind specific HLA class I or class II alleles. As appreciated by one of ordinary skill in the art, a large array of computer programming software and hardware options are available in the relevant art which can be employed to implement the motifs of the invention in order to evaluate (e.g., without limitation, to identify epitopes, identify epitope concentration per peptide length, or to generate analogs) known or unknown peptide sequences.

In accordance with the procedures described above, HCV peptide epitopes and analogs thereof that are able to bind HLA supertype groups or allele-specific HLA molecules have been identified (Tables VII-XX; Table XXII).

30 IV.H. Preparation of Peptide Epitopes

Peptides in accordance with the invention can be prepared synthetically, by recombinant DNA technology or chemical synthesis, or from natural sources such as native tumors or pathogenic organisms. Peptide epitopes may be synthesized individually or as polyepitopic peptides. Although the peptide will preferably be substantially free of

other naturally occurring host cell proteins and fragments thereof, in some embodiments the peptides may be synthetically conjugated to native fragments or particles.

The peptides in accordance with the invention can be a variety of lengths, and either in their neutral (uncharged) forms or in forms which are salts. The peptides in 5 accordance with the invention are either free of modifications such as glycosylation, side chain oxidation, or phosphorylation; or they contain these modifications, subject to the condition that modifications do not destroy the biological activity of the peptides as described herein.

The peptides of the invention can be prepared in a wide variety of ways. For the 10 preferred relatively short size, the peptides can be synthesized in solution or on a solid support in accordance with conventional techniques. Various automatic synthesizers are commercially available and can be used in accordance with known protocols. (See, for example, Stewart & Young, *SOLID PHASE PEPTIDE SYNTHESIS*, 2D. ED., Pierce Chemical Co., 1984). Further, individual peptide epitopes can be joined using chemical ligation to 15 produce larger peptides that are still within the bounds of the invention.

Alternatively, recombinant DNA technology can be employed wherein a nucleotide sequence which encodes an immunogenic peptide of interest is inserted into an expression vector, transformed or transfected into an appropriate host cell and cultivated under conditions suitable for expression. These procedures are generally known in the 20 art, as described generally in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989). Thus, recombinant polypeptides which comprise one or more peptide sequences of the invention can be used to present the appropriate T cell epitope.

The nucleotide coding sequence for peptide epitopes of the preferred lengths 25 contemplated herein can be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci, *et al.*, *J. Am. Chem. Soc.* 103:3185 (1981). Peptide analogs can be made simply by substituting the appropriate and desired nucleic acid base(s) for those that encode the native peptide sequence; exemplary nucleic acid substitutions are those that encode an amino acid defined by the motifs/supermotifs 30 herein. The coding sequence can then be provided with appropriate linkers and ligated into expression vectors commonly available in the art, and the vectors used to transform suitable hosts to produce the desired fusion protein. A number of such vectors and suitable host systems are now available. For expression of the fusion proteins, the coding sequence will be provided with operably linked start and stop codons, promoter and

terminator regions and usually a replication system to provide an expression vector for expression in the desired cellular host. For example, promoter sequences compatible with bacterial hosts are provided in plasmids containing convenient restriction sites for insertion of the desired coding sequence. The resulting expression vectors are 5 transformed into suitable bacterial hosts. Of course, yeast, insect or mammalian cell hosts may also be used, employing suitable vectors and control sequences.

It is often preferable that the peptide epitope be as small as possible while still maintaining substantially all of the immunologic activity of the native protein. When possible, it may be desirable to optimize HLA class I binding peptide epitopes of the 10 invention to a length of about 8 to about 13 amino acid residues, preferably 9 to 10. HLA class II binding peptide epitopes may be optimized to a length of about 6 to about 30 amino acids in length, preferably to between about 13 and about 20 residues. Preferably, the peptide epitopes are commensurate in size with endogenously processed pathogen-derived peptides or tumor cell peptides that are bound to the relevant HLA molecules, 15 however, the identification and preparation of peptides of other lengths can also be carried out using the techniques described herein.

In alternative embodiments, peptides of the invention can be linked as a polyepitopic peptide, or as a minigene that encodes a polyepitopic peptide.

In another embodiment, it is preferred to identify native peptide regions that 20 contain a high concentration of class I and/or class II epitopes. Such a sequence is generally selected on the basis that it contains the greatest number of epitopes per amino acid length. It is to be appreciated that epitopes can be present in a frame-shifted manner, e.g. a 10 amino acid long peptide could contain two 9 amino acid long epitopes and one 10 amino acid long epitope; upon intracellular processing, each epitope can be exposed 25 and bound by an HLA molecule upon administration of such a peptide. This larger, preferably multi-epitopic, peptide can be generated synthetically, recombinantly, or via cleavage from the native source.

IV.I. Assays to Detect T-Cell Responses

Once HLA binding peptides are identified, they can be tested for the ability to 30 elicit a T-cell response. The preparation and evaluation of motif-bearing peptides are described in PCT publications WO 94/20127 and WO 94/03205. Briefly, peptides comprising epitopes from a particular antigen are synthesized and tested for their ability to bind to the appropriate HLA proteins. These assays may involve evaluating the

binding of a peptide of the invention to purified HLA class I molecules in relation to the binding of a radioiodinated reference peptide. Alternatively, cells expressing empty class I molecules (*i.e.* lacking peptide therein) may be evaluated for peptide binding by immunofluorescent staining and flow microfluorimetry. Other assays that may be used to evaluate peptide binding include peptide-dependent class I assembly assays and/or the inhibition of CTL recognition by peptide competition. Those peptides that bind to the class I molecule, typically with an affinity of 500 nM or less, are further evaluated for their ability to serve as targets for CTLs derived from infected or immunized individuals, as well as for their capacity to induce primary *in vitro* or *in vivo* CTL responses that can give rise to CTL populations capable of reacting with selected target cells associated with a disease. Corresponding assays are used for evaluation of HLA class II binding peptides. HLA class II motif-bearing peptides that are shown to bind, typically at an affinity of 1000 nM or less, are further evaluated for the ability to stimulate CTL responses.

Conventional assays utilized to detect T cell responses include proliferation assays, lymphokine secretion assays, direct cytotoxicity assays, and limiting dilution assays. For example, antigen-presenting cells that have been incubated with a peptide can be assayed for the ability to induce CTL responses in responder cell populations. Antigen-presenting cells can be normal cells such as peripheral blood mononuclear cells or dendritic cells. Alternatively, mutant non-human mammalian cell lines that are deficient in their ability to load class I molecules with internally processed peptides and that have been transfected with the appropriate human class I gene, may be used to test for the capacity of the peptide to induce *in vitro* primary CTL responses.

Peripheral blood mononuclear cells (PBMCs) may be used as the responder cell source of CTL precursors. The appropriate antigen-presenting cells are incubated with peptide, after which the peptide-loaded antigen-presenting cells are then incubated with the responder cell population under optimized culture conditions. Positive CTL activation can be determined by assaying the culture for the presence of CTLs that kill radio-labeled target cells, both specific peptide-pulsed targets as well as target cells expressing endogenously processed forms of the antigen from which the peptide sequence was derived.

More recently, a method has been devised which allows direct quantification of antigen-specific T cells by staining with Fluorescein-labelled HLA tetrameric complexes (Altman, J. D. *et al.*, *Proc. Natl. Acad. Sci. USA* 90:10330, 1993; Altman, J. D. *et al.*, *Science* 274:94, 1996). Other relatively recent technical developments include staining

for intracellular lymphokines, and interferon release assays or ELISPOT assays.

Tetramer staining, intracellular lymphokine staining and ELISPOT assays all appear to be at least 10-fold more sensitive than more conventional assays (Lalvani, A. *et al.*, *J. Exp. Med.* 186:859, 1997; Dunbar, P. R. *et al.*, *Curr. Biol.* 8:413, 1998; Murali-Krishna, K. *et al.*, *Immunity* 8:177, 1998).

HTL activation may also be assessed using such techniques known to those in the art such as T cell proliferation and secretion of lymphokines, *e.g.* IL-2 (*see, e.g.* Alexander *et al.*, *Immunity* 1:751-761, 1994).

Alternatively, immunization of HLA transgenic mice can be used to determine immunogenicity of peptide epitopes. Several transgenic mouse models including mice with human A2.1, A11 (which can additionally be used to analyze HLA-A3 epitopes), and B7 alleles have been characterized and others (*e.g.*, transgenic mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed. Additional transgenic mouse models with other HLA alleles may be generated as necessary. Mice may be immunized with peptides emulsified in Incomplete Freund's Adjuvant and the resulting T cells tested for their capacity to recognize peptide-pulsed target cells and target cells transfected with appropriate genes. CTL responses may be analyzed using cytotoxicity assays described above. Similarly, HTL responses may be analyzed using such assays as T cell proliferation or secretion of lymphokines.

Exemplary immunogenic peptide epitopes are set out in Table XXIII.

IV.J. Use of Peptide Epitopes as Diagnostic Agents and for Evaluating Immune Responses

In one embodiment of the invention, HLA class I and class II binding peptides as described herein can be used as reagents to evaluate an immune response. The immune response to be evaluated can be induced by using as an immunogen any agent that may result in the production of antigen-specific CTLs or HTLs that recognize and bind to the peptide epitope(s) to be employed as the reagent. The peptide reagent need not be used as the immunogen. Assay systems that can be used for such an analysis include relatively recent technical developments such as tetramers, staining for intracellular lymphokines and interferon release assays, or ELISPOT assays.

For example, a peptide of the invention may be used in a tetramer staining assay to assess peripheral blood mononuclear cells for the presence of antigen-specific CTLs following exposure to a tumor cell antigen or an immunogen. The HLA-tetrameric

complex is used to directly visualize antigen-specific CTLs (see, e.g., Ogg *et al.*, *Science* 279:2103-2106, 1998; and Altman *et al.*, *Science* 174:94-96, 1996) and determine the frequency of the antigen-specific CTL population in a sample of peripheral blood mononuclear cells. A tetramer reagent using a peptide of the invention may be generated 5 as follows: A peptide that binds to an HLA molecule is refolded in the presence of the corresponding HLA heavy chain and β_2 -microglobulin to generate a trimolecular complex. The complex is biotinylated at the carboxyl terminal end of the heavy chain at a site that was previously engineered into the protein. Tetramer formation is then induced by the addition of streptavidin. By means of fluorescently labeled streptavidin, the 10 tetramer can be used to stain antigen-specific cells. The cells may then be identified, for example, by flow cytometry. Such an analysis may be used for diagnostic or prognostic purposes. Cells identified by the procedure can also be used for therapeutic purposes.

Peptides of the invention may also be used as reagents to evaluate immune recall responses. (see, e.g., Bertoni *et al.*, *J. Clin. Invest.* 100:503-513, 1997 and Penna *et al.*, *J. 15 Exp. Med.* 174:1565-1570, 1991.) For example, patient PBMC samples from individuals with HCV infection may be analyzed for the presence of antigen-specific CTLs or HTLs using specific peptides. A blood sample containing mononuclear cells may be evaluated by cultivating the PBMCs and stimulating the cells with a peptide of the invention. After an appropriate cultivation period, the expanded cell population may be analyzed, for 20 example, for cytotoxic activity (CTL) or for HTL activity.

The peptides may also be used as reagents to evaluate the efficacy of a vaccine. PBMCs obtained from a patient vaccinated with an immunogen may be analyzed using, for example, either of the methods described above. The patient is HLA typed, and peptide epitope reagents that recognize the allele-specific molecules present in that 25 patient are selected for the analysis. The immunogenicity of the vaccine is indicated by the presence of epitope-specific CTLs and/or HTLs in the PBMC sample.

The peptides of the invention may also be used to make antibodies, using techniques well known in the art (see, e.g. *CURRENT PROTOCOLS IN IMMUNOLOGY*, Wiley/Greene, NY; and *Antibodies A Laboratory Manual*, Harlow and Lane, Cold Spring 30 Harbor Laboratory Press, 1989), which may be useful as reagents to diagnose or monitor cancer. Such antibodies include those that recognize a peptide in the context of an HLA molecule, *i.e.*, antibodies that bind to a peptide-MHC complex.

IV.K. Vaccine Compositions

Vaccines and methods of preparing vaccines that contain an immunogenically effective amount of one or more peptides as described herein are further embodiments of the invention. Once appropriately immunogenic epitopes have been defined, they can be 5 sorted and delivered by various means, herein referred to as "vaccine" compositions. Such vaccine compositions can include, for example, lipopeptides (e.g., Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341, 1995), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres (see, e.g., Eldridge, *et al.*, *Molec. Immunol.* 28:287-294, 1991; Alonso *et al.*, *Vaccine* 12:299-306, 1994; Jones *et al.*, *Vaccine* 13:675-681, 1995), 10 peptide compositions contained in immune stimulating complexes (ISCOMS) (see, e.g., Takahashi *et al.*, *Nature* 344:873-875, 1990; Hu *et al.*, *Clin Exp Immunol.* 113:235-243, 1998), multiple antigen peptide systems (MAPs) (see e.g., Tam, J. P., *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413, 1988; Tam, J.P., *J. Immunol. Methods* 196:17-32, 1996), viral delivery vectors (Perkus, M. E. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. 15 H. E., ed., p. 379, 1996; Chakrabarti, S. *et al.*, *Nature* 320:535, 1986; Hu, S. L. *et al.*, *Nature* 320:537, 1986; Kieny, M.-P. *et al.*, *AIDS Bio/Technology* 4:790, 1986; Top, F. H. *et al.*, *J. Infect. Dis.* 124:148, 1971; Chanda, P. K. *et al.*, *Virology* 175:535, 1990), particles of viral or synthetic origin (e.g., Kofler, N. *et al.*, *J. Immunol. Methods*. 192:25, 1996; Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993; Falo, L. D., Jr. *et al.*, *Nature 20 Med.* 7:649, 1995), adjuvants (Warren, H. S., Vogel, F. R., and Chedid, L. A. *Annu. Rev. Immunol.* 4:369, 1986; Gupta, R. K. *et al.*, *Vaccine* 11:293, 1993), liposomes (Reddy, R. *et al.*, *J. Immunol.* 148:1585, 1992; Rock, K. L., *Immunol. Today* 17:131, 1996), or, 25 naked or particle absorbed cDNA (Ulmer, J. B. *et al.*, *Science* 259:1745, 1993; Robinson, H. L., Hunt, L. A., and Webster, R. G., *Vaccine* 11:957, 1993; Shiver, J. W. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., *Annu. Rev. Immunol.* 12:923, 1994 and Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993). Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used. 30 Vaccines of the invention include nucleic acid-mediated modalities. DNA or RNA encoding one or more of the peptides of the invention can also be administered to a patient. This approach is described, for instance, in Wolff *et. al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based

delivery technologies include "naked DNA", facilitated (bupivacaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (see, e.g., U.S. Patent No. 5,922,687).

For therapeutic or prophylactic immunization purposes, the peptides of the invention can also be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. As an example of this approach, vaccinea virus is used as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into a host bearing a tumor, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits a host 10 CTL and/or HTL response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, e.g. adeno and adeno-associated virus 15 vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein.

Furthermore, vaccines in accordance with the invention encompass compositions of one or more of the claimed peptide(s). A peptide can be present in a vaccine individually. Alternatively, the peptide can exist as a homopolymer comprising 20 multiple copies of the same peptide, or as a heteropolymer of various peptides. Polymers have the advantage of increased immunological reaction and, where different peptide epitopes are used to make up the polymer, the additional ability to induce antibodies and/or CTLs that react with different antigenic determinants of the pathogenic organism or tumor-related peptide targeted for an immune response. The composition can be a 25 naturally occurring region of an antigen or can be prepared, e.g., recombinantly or by chemical synthesis.

Carriers that can be used with vaccines of the invention are well known in the art, and include, e.g., thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza, hepatitis B virus 30 core protein, and the like. The vaccines can contain a physiologically tolerable (i.e., acceptable) diluent such as water, or saline, preferably phosphate buffered saline. The vaccines also typically include an adjuvant. Adjuvants such as incomplete Freund's adjuvant, aluminum phosphate, aluminum hydroxide, or alum are examples of materials well known in the art. Additionally, as disclosed herein, CTL responses can be primed by

conjugating peptides of the invention to lipids, such as tripalmitoyl-S-glycerylcysteinylseryl-serine (P₃CSS).

Upon immunization with a peptide composition in accordance with the invention, via injection, aerosol, oral, transdermal, transmucosal, intrapleural, intrathecal, or other suitable routes, the immune system of the host responds to the vaccine by producing large amounts of CTLs and/or HTLs specific for the desired antigen. Consequently, the host becomes at least partially immune to later infection, or at least partially resistant to developing an ongoing chronic infection, or derives at least some therapeutic benefit when the antigen was tumor-associated.

10 In some embodiments it may be desirable to combine the class I peptide components with components that induce or facilitate neutralizing antibody responses to the target antigen of interest, particularly to viral envelope antigens. A preferred embodiment of such a composition comprises class I and class II epitopes in accordance with the invention. An alternative embodiment of such a composition comprises a class I and/or class II epitope in accordance with the invention, along with a PADRE™ (Epimmune, San Diego, CA) molecule (described, for example, in U.S. Patent Number 5,736,142).

15

A vaccine of the invention can also include antigen-presenting cells, such as dendritic cells, as a vehicle to present peptides of the invention. Vaccine compositions can be created *in vitro*, following dendritic cell mobilization and harvesting, whereby loading of dendritic cells occurs *in vitro*. For example, dendritic cells are transfected, *e.g.*, with a minigene in accordance with the invention. The dendritic cell can then be administered to a patient to elicit immune responses *in vivo*.

20 Antigenic peptides are used to elicit a CTL and/or HTL response *ex vivo*, as well. The resulting CTL or HTL cells, can be used to treat tumors in patients that do not respond to other conventional forms of therapy, or will not respond to a therapeutic vaccine peptide or nucleic acid in accordance with the invention. *Ex vivo* CTL or HTL responses to a particular tumor-associated antigen are induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of antigen-presenting cells (APC), such as dendritic cells, and the appropriate immunogenic peptide. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction

25

30

(HTL) of their specific target cell (an infected cell or a tumor cell). Transfected dendritic cells may also be used as antigen presenting cells.

The vaccine compositions of the invention can also be used in combination with antiviral drugs such as interferon- α , or other treatments for viral infection.

5 Preferably, the following principles are utilized when selecting an array of epitopes for inclusion in a polyepitopic composition for use in a vaccine, or for selecting discrete epitopes to be included in a vaccine and/or to be encoded by nucleic acids such as a minigene. It is preferred that each of the following principles are balanced in order to make the selection. The multiple epitopes to be incorporated in a given vaccine
10 composition may be, but need not be, contiguous in sequence in the native antigen from which the epitopes are derived.

15 Preferably, the following principles are utilized when selecting an array of epitopes for inclusion in a polyepitopic composition for use in a vaccine, or for selecting discrete epitopes to be included in a vaccine and/or to be encoded by nucleic acids such as a minigene. Exemplary epitopes that may be utilized in a vaccine to treat or prevent HCV infection are set out in Tables XXVI-XXIX, and Table XXXII. It is preferred that each of the following principles are balanced in order to make the selection.

20 1.) Epitopes are selected which, upon administration, mimic immune responses that have been observed to be correlated with HCV clearance. For HLA Class I this includes 3-4 epitopes that come from at least one antigen of HCV. For HLA Class II a similar rationale is employed; again 3-4 epitopes are selected from at least one HCV antigen (see e.g., Rosenberg *et al.*, *Science* 278:1447-1450).

25 2.) Epitopes are selected that have the requisite binding affinity established to be correlated with immunogenicity: for HLA Class I an IC_{50} of 500 nM or less, or for Class II an IC_{50} of 1000 nM or less.

30 3.) Sufficient supermotif bearing-peptides, or a sufficient array of allele-specific motif-bearing peptides, are selected to give broad population coverage. For example, it is preferable to have at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess the breadth, or redundancy of, population coverage.

4.) When selecting epitopes from cancer-related antigens it is often preferred to select analogs because the patient may have developed tolerance to the native epitope.

When selecting epitopes for infectious disease-related antigens it is preferable to select either native or analoged epitopes.

5.) Of particular relevance are epitopes referred to as "nested epitopes."

Nested epitopes occur where at least two epitopes overlap in a given peptide sequence. A

5 nested peptide sequence can comprise both HLA class I and HLA class II epitopes.

When providing nested epitopes, it is preferable to provide a sequence that has the greatest number of epitopes per provided sequence. Preferably, one avoids providing a peptide that is any longer than the amino terminus of the amino terminal epitope and the carboxyl terminus of the carboxyl terminal epitope in the peptide. When providing a

10 longer peptide sequence, such as a sequence comprising nested epitopes, it is important to screen the sequence in order to insure that it does not have pathological or other deleterious biological properties.

6.) If a polyepitopic protein is created, or when creating a minigene, an objective is to generate the smallest peptide that encompasses the epitopes of interest.

15 This principle is similar, if not the same as that employed when selecting a peptide comprising nested epitopes. However, with an artificial polyepitopic peptide, the size minimization objective is balanced against the need to integrate any spacer sequences between epitopes in the polyepitopic protein. Spacer amino acid residues can be introduced to avoid junctional epitopes (an epitope recognized by the immune system, not 20 present in the target antigen, and only created by the man-made juxtaposition of epitopes), or to facilitate cleavage between epitopes and thereby enhance epitope presentation. Junctional epitopes are generally to be avoided because the recipient may generate an immune response to that non-native epitope. Of particular concern is a junctional epitope that is a "dominant epitope." A dominant epitope may lead to such a 25 zealous response that immune responses to other epitopes are diminished or suppressed.

Examples of polyepitopic vaccine compositions designed based on the above criteria can include epitopes from the core, S, E1, NS1/E2, NS2, NS3, NS4, and NS5 domains of the HCV polyprotein. These regions encompass the following amino acid sequences using numbering relative to the prototype HCV-1 strain (Genbank accession 30 number M62321; *see, e.g.*, US Patent Nos. 5,683,864 and 5,670,153): C domain (amino acids 1-120); S (amino acids 120-400); NS3 (amino acids 1050-1640); NS4 (amino acids 1640-2000); NS5 (amino acids 2000-3011); and envelop proteins, E1 and E2/NS1, encompassing amino acids 192-750. Amino acids 750 to 1050 are designated as domain X as applied to the present invention. As appreciated by one of ordinary skill in the art,

the designation of the amino acid range for each domain may diverge to some extent from that of HCV-1 depending on the strain of HCV. One of ordinary skill in the art, when looking at an HCV polyprotein sequence, would readily be able to determine the domain boundaries.

5 Specific embodiments of the polyepitopic compositions of the present invention include a pharmaceutical composition comprising a pharmaceutically acceptable carrier and combination of motif-bearing peptides that are immunologically cross-reactive with peptides of HCV-1, wherein at least one of the peptides bears a motif of Table Ia, and further wherein the combination of motif-bearing peptides consists of: a) one or more

10 peptides comprising at least 8 amino acids from an HCV C domain; b) one or more peptides comprising at least 8 amino acids of a further domain selected from the group consisting of: an S domain, an NS3 domain, an NS4 domain, or an NS5 domain, and; c) optionally, one or more motif-bearing peptides from one or more additional HCV domains with a *proviso* that an additional domain is not a further domain listed in "b".

15 Preferably, such a pharmaceutical composition may additionally comprise one or more distinct HCV motif-bearing peptide(s) comprising at least 8 amino acids of an X domain or, alternatively, the composition may further comprise additional HCV motif-bearing peptide(s) that are from an envelope domain, the envelope domain peptide(s) consisting of one or more copies of a single HCV peptide comprising at least 8 amino acids of an

20 envelope domain.

 In another embodiment, the polyepitopic pharmaceutical composition may comprise a pharmaceutically acceptable carrier and combination of motif-bearing peptides that are immunologically cross-reactive with HCV-1 peptides, the peptides from multiple domains of HCV, wherein at least one of the peptides bears a motif of Table Ia, and wherein the combination of motif-bearing peptides consists essentially of: a) one or more peptides comprising at least 8 amino acids from a C domain; and, b) one or more peptides comprising at least 8 amino acids from an S, NS3, NS4, or NS5 domain, and, one HCV peptide comprising at least 8 amino acids of an envelope domain. Such a composition may further comprise one or more HCV motif-bearing peptides comprising at least 8 amino acids of an X domain.

 Alternatively, a pharmaceutical composition of the invention may comprise: a) a pharmaceutically acceptable carrier; and, b) a combination of one or more motif-bearing peptides of at least 8 amino acids derived from one or more hepatitis C virus (HCV) domains, wherein said peptides are cross-reactive with peptides of HCV-1, with a *proviso*

that the combination does not include a peptide of at least 8 amino acids from an HCV C domain, and wherein at least one of the peptides bears a motif of Table Ia, said domains selected from the group consisting of: an S domain; an NS3 domain; an NS4 domain; an NS5 domain; and, an X domain. Such a composition may additionally comprise motif-
5 bearing HCV envelope peptide(s) consisting of one or more copies of a single HCV peptide comprising at least 8 amino acids of an envelope domain.

Lastly, an embodiment of the invention may comprise a pharmaceutical composition comprising a pharmaceutically acceptable carrier and combination of two or more motif-bearing peptides from a single domain of an HCV-1 strain, said peptides
10 immunologically cross-reactive with peptides of an HCV-1 antigen, wherein at least one of the peptides bears a motif of Table Ia, and the peptides are derived from HCV, and the HCV domain is selected from the group consisting of: a C domain; an S domain; an NS3 domain; an NS4 domain; an NS5 domain; an X domain; or, an envelope domain from a single HCV strain, with a *proviso* that the envelope domain is other than a variable
15 envelope domain.

In the embodiments set forth, “peptides immunologically cross-reactive with HCV-1” refers to peptides that are bound by the same antibody; “derived from” refers to a fragment or subsequence and conservatively modified variants thereof.

20 IV.K.1. Minigene Vaccines

A number of different approaches are available which allow simultaneous delivery of multiple epitopes. Nucleic acids encoding the peptides of the invention are a particularly useful embodiment of the invention. Epitopes for inclusion in a minigene are preferably selected according to the guidelines set forth in the previous section. A
25 preferred means of administering nucleic acids encoding the peptides of the invention uses minigene constructs encoding a peptide comprising one or multiple epitopes of the invention.

The use of multi-epitope minigenes is described below and in, *e.g.*, co-pending application U.S.S.N. 09/311,784; An, L. and Whitton, J. L., *J. Virol.* 71:2292, 1997;
30 Thomson, S. A. *et al.*, *J. Immunol.* 157:822, 1996; Whitton, J. L. *et al.*, *J. Virol.* 67:348, 1993; Hanke, R. *et al.*, *Vaccine* 16:426, 1998. For example, a multi-epitope DNA plasmid encoding supermotif- and/or motif-bearing HCV epitopes derived from multiple regions of the HCV polyprotein sequence, the PADRE™ universal helper T cell epitope (or

multiple HTL epitopes from HCV), and an endoplasmic reticulum-translocating signal sequence can be engineered.

The immunogenicity of a multi-epitopic minigene can be tested in transgenic mice to evaluate the magnitude of CTL induction responses against the epitopes tested.

5 Further, the immunogenicity of DNA-encoded epitopes *in vivo* can be correlated with the *in vitro* responses of specific CTL lines against target cells transfected with the DNA plasmid. Thus, these experiments can show that the minigene serves to both: 1.) generate a CTL response and 2.) that the induced CTLs recognized cells expressing the encoded epitopes.

10 For example, to create a DNA sequence encoding the selected epitopes (minigene) for expression in human cells, the amino acid sequences of the epitopes may be reverse translated. A human codon usage table can be used to guide the codon choice for each amino acid. These epitope-encoding DNA sequences may be directly adjoined, so that when translated, a continuous polypeptide sequence is created. To optimize expression 15 and/or immunogenicity, additional elements can be incorporated into the minigene design. Examples of amino acid sequences that can be reverse translated and included in the minigene sequence include: HLA class I epitopes, HLA class II epitopes, a ubiquitination signal sequence, and/or an endoplasmic reticulum targeting signal. In addition, HLA presentation of CTL and HTL epitopes may be improved by including 20 synthetic (*e.g.* poly-alanine) or naturally-occurring flanking sequences adjacent to the CTL or HTL epitopes; these larger peptides comprising the epitope(s) are within the scope of the invention.

The minigene sequence may be converted to DNA by assembling oligonucleotides that encode the plus and minus strands of the minigene. Overlapping oligonucleotides 25 (30-100 bases long) may be synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides can be joined, for example, using T4 DNA ligase. This synthetic minigene, encoding the epitope polypeptide, can then be cloned into a desired expression vector.

Standard regulatory sequences well known to those of skill in the art are 30 preferably included in the vector to ensure expression in the target cells. Several vector elements are desirable: a promoter with a down-stream cloning site for minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (*e.g.* ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, *e.g.*, the human cytomegalovirus

(hCMV) promoter. See, *e.g.*, U.S. Patent Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences.

Additional vector modifications may be desired to optimize minigene expression and immunogenicity. In some cases, introns are required for efficient gene expression, 5 and one or more synthetic or naturally-occurring introns could be incorporated into the transcribed region of the minigene. The inclusion of mRNA stabilization sequences and sequences for replication in mammalian cells may also be considered for increasing minigene expression.

Once an expression vector is selected, the minigene is cloned into the polylinker 10 region downstream of the promoter. This plasmid is transformed into an appropriate *E. coli* strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the minigene, as well as all other elements included in the vector, are confirmed using restriction mapping and DNA sequence analysis. Bacterial cells harboring the correct plasmid can be stored as a master cell bank and a working cell bank.

15 In addition, immunostimulatory sequences (ISSs or CpGs) appear to play a role in the immunogenicity of DNA vaccines. These sequences may be included in the vector, outside the minigene coding sequence, if desired to enhance immunogenicity.

In some embodiments, a bi-cistronic expression vector which allows production of 20 both the minigene-encoded epitopes and a second protein (included to enhance or decrease immunogenicity) can be used. Examples of proteins or polypeptides that could beneficially enhance the immune response if co-expressed include cytokines (*e.g.*, IL-2, IL-12, GM-CSF), cytokine-inducing molecules (*e.g.*, LeIF), costimulatory molecules, or for HTL responses, pan-DR binding proteins (PADRE™, Epimmune, San Diego, CA). Helper (HTL) epitopes can be joined to intracellular targeting signals and expressed 25 separately from expressed CTL epitopes; this allows direction of the HTL epitopes to a cell compartment different than that of the CTL epitopes. If required, this could facilitate more efficient entry of HTL epitopes into the HLA class II pathway, thereby improving HTL induction. In contrast to HTL or CTL induction, specifically decreasing the immune response by co-expression of immunosuppressive molecules (*e.g.* TGF- β) may be 30 beneficial in certain diseases.

Therapeutic quantities of plasmid DNA can be produced for example, by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are used to inoculate growth medium, and grown to saturation in shaker flasks or a bioreactor

according to well known techniques. Plasmid DNA can be purified using standard bioseparation technologies such as solid phase anion-exchange resins supplied by QIAGEN, Inc. (Valencia, California). If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

5 Purified plasmid DNA can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized DNA in sterile phosphate-buffer saline (PBS). This approach, known as "naked DNA," is currently being used for intramuscular (IM) administration in clinical trials. To maximize the immunotherapeutic effects of minigene DNA vaccines, an alternative method for

10 formulating purified plasmid DNA may be desirable. A variety of methods have been described, and new techniques may become available. Cationic lipids can also be used in the formulation (see, *e.g.*, as described by WO 93/24640; Mannino & Gould-Fogerite, *BioTechniques* 6(7): 682 (1988); U.S. Pat No. 5,279,833; WO 91/06309; and Felgner, *et al.*, *Proc. Nat'l Acad. Sci. USA* 84:7413 (1987). In addition, glycolipids, fusogenic

15 liposomes, peptides and compounds referred to collectively as protective, interactive, non-condensing compounds (PINC) could also be complexed to purified plasmid DNA to influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

Target cell sensitization can be used as a functional assay for expression and HLA class I presentation of minigene-encoded CTL epitopes. For example, the plasmid DNA is introduced into a mammalian cell line that is suitable as a target for standard CTL chromium release assays. The transfection method used will be dependent on the final formulation. Electroporation can be used for "naked" DNA, whereas cationic lipids allow direct *in vitro* transfection. A plasmid expressing green fluorescent protein (GFP) can be co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). These cells are then chromium-51 (^{51}Cr) labeled and used as target cells for epitope-specific CTL lines; cytolysis, detected by ^{51}Cr release, indicates both production of, and HLA presentation of, minigene-encoded CTL epitopes. Expression of HTL epitopes may be evaluated in an analogous manner using assays to assess HTL activity.

In vivo immunogenicity is a second approach for functional testing of minigene DNA formulations. Transgenic mice expressing appropriate human HLA proteins are immunized with the DNA product. The dose and route of administration are formulation dependent (*e.g.*, IM for DNA in PBS, intraperitoneal (IP) for lipid-complexed DNA).

Twenty-one days after immunization, splenocytes are harvested and restimulated for 1 week in the presence of peptides encoding each epitope being tested. Thereafter, for CTL effector cells, assays are conducted for cytolysis of peptide-loaded, ^{51}Cr -labeled target cells using standard techniques. Lysis of target cells that were sensitized by HLA loaded 5 with peptide epitopes, corresponding to minigene-encoded epitopes, demonstrates DNA vaccine function for *in vivo* induction of CTLs. Immunogenicity of HTL epitopes is evaluated in transgenic mice in an analogous manner.

Alternatively, the nucleic acids can be administered using ballistic delivery as described, for instance, in U.S. Patent No. 5,204,253. Using this technique, particles 10 comprised solely of DNA are administered. In a further alternative embodiment, DNA can be adhered to particles, such as gold particles.

IV.K.2. Combinations of CTL Peptides with Helper Peptides

Vaccine compositions comprising the peptides of the present invention, or analogs 15 thereof, which have immunostimulatory activity may be modified to provide desired attributes, such as improved serum half life, or to enhance immunogenicity.

For instance, the ability of the peptides to induce CTL activity can be enhanced by linking the peptide to a sequence which contains at least one epitope that is capable of inducing a T helper cell response. The use of T helper epitopes in conjunction with CTL 20 epitopes to enhance immunogenicity is illustrated, for example, in co-pending U.S.S.N. 08/820360, U.S.S.N. 08/197,484, and U.S.S.N. 08/464,234.

Particularly preferred CTL epitope/HTL epitope conjugates are linked by a spacer molecule. The spacer is typically comprised of relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under 25 physiological conditions. The spacers are typically selected from, *e.g.*, Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer need not be comprised of the same residues and thus may be a hetero- or homo-oligomer. When present, the spacer will usually be at least one or two residues, more usually three to six residues. Alternatively, the CTL 30 peptide may be linked to the T helper peptide without a spacer.

Although the CTL peptide epitope can be linked directly to the T helper peptide epitope, often CTL epitope/HTL epitope conjugates are linked by a spacer molecule. The spacer is typically comprised of relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological

conditions. The spacers are typically selected from, e.g., Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer need not be comprised of the same residues and thus may be a hetero- or homo-oligomer. When present, the spacer will usually be at least one or two 5 residues, more usually three to six residues. The CTL peptide epitope can be linked to the T helper peptide epitope either directly or via a spacer either at the amino or carboxy terminus of the CTL peptide. The amino terminus of either the immunogenic peptide or the T helper peptide may be acylated.

HTL peptide epitopes can also be modified to alter their biological properties. For 10 example, peptides comprising HTL epitopes can contain D-amino acids to increase their resistance to proteases and thus extend their serum half-life. Also, the epitope peptides of the invention can be conjugated to other molecules such as lipids, proteins or sugars, or any other synthetic compounds, to increase their biological activity. Specifically, the T helper peptide can be conjugated to one or more palmitic acid chains at either the amino 15 or carboxyl termini.

In certain embodiments, the T helper peptide is one that is recognized by T helper cells present in the majority of the population. This can be accomplished by selecting amino acid sequences that bind to many, most, or all of the HLA class II molecules. These are known as "loosely HLA-restricted" or "promiscuous" T helper sequences. 20 Examples of amino acid sequences that are promiscuous include sequences from antigens such as tetanus toxoid at positions 830-843 (QYIKANSKFIGITE), *Plasmodium falciparum* CS protein at positions 378-398 (DIEKKIAKMEKASSVFNVVNS), and *Streptococcus* 18kD protein at positions 116 (GAVDSILGGVATYGAA). Other examples include peptides bearing a DR 1-4-7 supermotif, or either of the DR3 motifs.

25 Alternatively, it is possible to prepare synthetic peptides capable of stimulating T helper lymphocytes, in a loosely HLA-restricted fashion, using amino acid sequences not found in nature (see, e.g., PCT publication WO 95/07707). These synthetic compounds called Pan-DR-binding epitopes (e.g., PADRE™, Epimmune, Inc., San Diego, CA) are designed to most preferably bind most HLA-DR (human HLA class II) molecules. For 30 instance, a pan-DR-binding epitope peptide having the formula: aKXVVWANTLKAa, where "X" is either cyclohexylalanine, phenylalanine, or tyrosine, and a is either D-alanine or L-alanine, has been found to bind to most HLA-DR alleles, and to stimulate the response of T helper lymphocytes from most individuals, regardless of their HLA type.

An alternative of a pan-DR binding epitope comprises all "L" natural amino acids and can be provided in the form of nucleic acids that encode the epitope.

In some embodiments it may be desirable to include in the pharmaceutical compositions of the invention at least one component which primes cytotoxic T

- 5 lymphocytes. Lipids have been identified as agents capable of priming CTL *in vivo* against viral antigens. For example, palmitic acid residues can be attached to the ϵ -and α -amino groups of a lysine residue and then linked, *e.g.*, via one or more linking residues such as Gly, Gly-Gly-, Ser, Ser-Ser, or the like, to an immunogenic peptide. The lipidated peptide can then be administered either directly in a micelle or particle, 10 incorporated into a liposome, or emulsified in an adjuvant, *e.g.*, incomplete Freund's adjuvant. In a preferred embodiment, a particularly effective immunogenic comprises palmitic acid attached to ϵ - and α - amino groups of Lys, which is attached via linkage, *e.g.*, Ser-Ser, to the amino terminus of the immunogenic peptide.

As another example of lipid priming of CTL responses, *E. coli* lipoproteins, such 15 as tripalmitoyl-S-glycerylcysteinylseryl- serine (P₃CSS) can be used to prime virus specific CTL when covalently attached to an appropriate peptide. (*See, e.g.*, Deres, *et al.*, *Nature* 342:561, 1989). Peptides of the invention can be coupled to P₃CSS, for example, and the lipopeptide administered to an individual to specifically prime a CTL response to the target antigen. Moreover, because the induction of neutralizing antibodies can also be 20 primed with P₃CSS-conjugated epitopes, two such compositions can be combined to more effectively elicit both humoral and cell-mediated responses to infection.

As noted herein, additional amino acids can be added to the termini of a peptide to provide for ease of linking peptides one to another, for coupling to a carrier support or larger peptide, for modifying the physical or chemical properties of the peptide or 25 oligopeptide, or the like. Amino acids such as tyrosine, cysteine, lysine, glutamic or aspartic acid, or the like, can be introduced at the C- or N-terminus of the peptide or oligopeptide, particularly class I peptides. However, it is to be noted that modification at the carboxyl terminus of a CTL epitope may, in some cases, alter binding characteristics of the peptide. In addition, the peptide or oligopeptide sequences can differ from the 30 natural sequence by being modified by terminal-NH₂ acylation, *e.g.*, by alkanoyl (C₁-C₂₀) or thioglycolyl acetylation, terminal-carboxyl amidation, *e.g.*, ammonia, methylamine, *etc.* In some instances these modifications may provide sites for linking to a support or other molecule.

Vaccine Compositions Comprising Dendritic Cells Pulsed with CTL and/or HTL Peptides

An embodiment of a vaccine composition in accordance with the invention comprises *ex vivo* administration of a cocktail of epitope-bearing peptides to PBMC, or 5 isolated DC therefrom, from the patient's blood. A pharmaceutical to facilitate harvesting of DC can be used, such as GM-CSF/IL-4. After pulsing the DC with peptides and prior to reinfusion into patients, the DC are washed to remove unbound peptides. In this embodiment, a vaccine comprises peptide-pulsed DCs which present the pulsed peptide epitopes complexed with HLA molecules on their surfaces. The vaccine is then 10 administered to the patient.

IV.L. Administration of Vaccines for Therapeutic or Prophylactic Purposes

The peptides of the present invention and pharmaceutical and vaccine compositions of the invention are useful for administration to mammals, particularly 15 humans, to treat and/or prevent HCV infection. Vaccine compositions containing the peptides of the invention are administered to a patient infected with HCV or to an individual susceptible to, or otherwise at risk for, HCV infection to elicit an immune response against HCV antigens and thus enhance the patient's own immune response capabilities. In therapeutic applications, peptide and/or nucleic acid compositions are 20 administered to a patient in an amount sufficient to elicit an effective CTL and/or HTL response to the virus antigen and to cure or at least partially arrest or slow symptoms and/or complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, *e.g.*, the particular composition administered, the manner of administration, the stage and severity 25 of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician.

The vaccine compositions of the invention may also be used purely as prophylactic agents. Generally the dosage for an initial prophylactic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 30 1000 μ g and the higher value is about 10,000; 20,000; 30,000; or 50,000 μ g. Dosage values for a human typically range from about 500 μ g to about 50,000 μ g per 70 kilogram patient. This is followed by boosting dosages of between about 1.0 μ g to about 50,000 μ g of peptide administered at defined intervals from about four weeks to six months after the

initial administration of vaccine. The immunogenicity of the vaccine may be assessed by measuring the specific activity of CTL and HTL obtained from a sample of the patient's blood.

As noted above, peptides comprising CTL and/or HTL epitopes of the invention 5 induce immune responses when presented by HLA molecules and contacted with a CTL or HTL specific for an epitope comprised by the peptide. The manner in which the peptide is contacted with the CTL or HTL is not critical to the invention. For instance, the peptide can be contacted with the CTL or HTL either *in vivo* or *in vitro*. If the contacting occurs *in vivo*, the peptide itself can be administered to the patient, or other 10 vehicles, *e.g.*, DNA vectors encoding one or more peptides, viral vectors encoding the peptide(s), liposomes and the like, can be used, as described herein. When the peptide is contacted *in vitro*, the vaccinating agent can comprise a population of cells, *e.g.*, peptide-pulsed dendritic cells, or TAA-specific CTLs, which have been induced by pulsing antigen-presenting cells *in vitro* with the peptide. Such a cell population is subsequently 15 administered to a patient in a therapeutically effective dose.

The peptides or DNA encoding them can be administered individually or as fusions of one or more peptide sequences.

For pharmaceutical compositions, the immunogenic peptides of the invention, or 20 DNA encoding them, are generally administered to an individual already infected with HCV. The peptides or DNA encoding them can be administered individually or as fusions of one or more peptide sequences. Those in the incubation phase or the acute phase of infection can be treated with the immunogenic peptides separately or in conjunction with other treatments, as appropriate.

For therapeutic use, administration should generally begin at the first diagnosis of 25 HCV infection. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. In chronic infection, loading doses followed by boosting doses may be required.

Treatment of an infected individual with the compositions of the invention may 30 hasten resolution of the infection in acutely infected individuals. For those individuals susceptible (or predisposed) to developing chronic infection, the compositions are particularly useful in methods for preventing the evolution from acute to chronic infection. Where susceptible individuals are identified prior to or during infection, the composition can be targeted to them, thus minimizing the need for administration to a larger population.

The peptide or other compositions used for the treatment or prophylaxis of HCV infection can be used, *e.g.*, in persons who have not manifested symptoms of disease but who act as a disease vector. In this context, it is generally important to provide an amount of the peptide epitope delivered by a mode of administration sufficient to 5 effectively stimulate a cytotoxic T cell response; compositions which stimulate helper T cell responses can also be given in accordance with this embodiment of the invention.

The dosage for an initial therapeutic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1000 μ g and the higher value is about 10,000; 20,000; 30,000; or 50,000 μ g. Dosage values for a human 10 typically range from about 500 μ g to about 50,000 μ g per 70 kilogram patient. Boosting dosages of between about 1.0 μ g to about 50000 μ g of peptide pursuant to a boosting regimen over weeks to months may be administered depending upon the patient's response and condition as determined by measuring the specific activity of CTL and HTL obtained from the patient's blood. The peptides and compositions of the present 15 invention may be employed in serious disease states, that is, life-threatening or potentially life threatening situations. In such cases, as a result of the minimal amounts of extraneous substances and the relative nontoxic nature of the peptides in preferred compositions of the invention, it is possible and may be felt desirable by the treating physician to administer substantial excesses of these peptide compositions relative to 20 these stated dosage amounts.

Thus, for treatment of chronic infection, a representative dose is in the range disclosed above, namely where the lower value is about 1, 5, 50, 500, or 1000 μ g and the higher value is about 10,000; 20,000; 30,000; or 50,000 μ g, preferably from about 500 μ g to about 50,000 μ g per 70 kilogram patient. Initial doses followed by boosting doses at 25 established intervals, *e.g.*, from four weeks to six months, may be required, possibly for a prolonged period of time to effectively immunize an individual. In the case of chronic infection, administration should continue until at least clinical symptoms or laboratory tests indicate that the viral infection has been eliminated or substantially abated and for a period thereafter. The dosages, routes of administration, and dose schedules are adjusted 30 in accordance with methodologies known in the art.

The pharmaceutical compositions for therapeutic treatment are intended for parenteral, topical, oral, intrathecal, or local administration. Preferably, the pharmaceutical compositions are administered parentally, *e.g.*, intravenously,

subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic peptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, *e.g.*, water, buffered water, 0.8% saline, 0.3% 5 glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain pharmaceutically acceptable auxiliary substances as required 10 to approximate physiological conditions, such as pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservatives, and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, *etc.*

15 The concentration of peptides of the invention in the pharmaceutical formulations can vary widely, *i.e.*, from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, *etc.*, in accordance with the particular mode of administration selected.

20 A human unit dose form of the peptide composition is typically included in a pharmaceutical composition that comprises a human unit dose of an acceptable carrier, preferably an aqueous carrier, and is administered in a volume of fluid that is known by those of skill in the art to be used for administration of such compositions to humans (*see, e.g., Remington's Pharmaceutical Sciences*, 17th Edition, A. Gennaro, Editor, Mack Publishing Co., Easton, Pennsylvania, 1985).

25 The peptides of the invention may also be administered via liposomes, which serve to target the peptides to a particular tissue, such as lymphoid tissue, or to target selectively to infected cells, as well as to increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations, the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a 30 molecule which binds to a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the peptide compositions. Liposomes for use in accordance with the invention are formed

from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, *e.g.*, liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing

5 liposomes, as described in, *e.g.*, Szoka, *et al.*, *Ann. Rev. Biophys. Bioeng.* 9:467 (1980), and U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

For targeting cells of the immune system, a ligand to be incorporated into the liposome can include, *e.g.*, antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a 10 peptide may be administered intravenously, locally, topically, *etc.* in a dose which varies according to, *inter alia*, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium 15 stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

20 For aerosol administration, the immunogenic peptides are preferably supplied in finely divided form along with a surfactant and propellant. Typical percentages of peptides are 0.01%-20% by weight, preferably 1%-10%. The surfactant must, of course, be nontoxic, and preferably soluble in the propellant. Representative of such agents are the esters or partial esters of fatty acids containing from 6 to 22 carbon atoms, such as 25 caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed esters, such as mixed or natural glycerides may be employed. The surfactant may constitute 0.1%-20% by weight of the composition, preferably 0.25-5%. The balance of the composition is ordinarily propellant. A carrier can also be included, as desired, as with, *e.g.*, lecithin for intranasal 30 delivery.

IV.M. Kits

The peptide and nucleic acid compositions of this invention can be provided in kit form together with instructions for vaccine administration. Typically the kit would

include desired peptide compositions in a container, preferably in unit dosage form and instructions for administration. An alternative kit would include a minigene construct with desired nucleic acids of the invention in a container, preferably in unit dosage form together with instructions for administration. Lymphokines such as IL-2 or IL-12 may 5 also be included in the kit. Other kit components that may also be desirable include, for example, a sterile syringe, booster dosages, and other desired excipients.

The invention will be described in greater detail by way of specific examples. The following examples are offered for illustrative purposes, and are not intended to limit the invention in any manner. Those of skill in the art will readily recognize a variety of 10 non-critical parameters that can be changed or modified to yield alternative embodiments in accordance with the invention.

V. EXAMPLES

As in many viral diseases, there is evidence that clearance of HCV is mediated by 15 CTL. In a study of primary HCV infection in six chimpanzees, four progressed to chronic infection (Cooper *et al.*, abstract, 19th US-Japan Hepatitis Joint Panel Meeting, January 27-29, 1998). It was found that these four animals showed either no CTL response or a very narrowly focused response during early infection. In contrast, in the remaining two animals that resolved the infection, a broad CTL response was observed 20 against multiple HCV proteins, some of which were conserved. Weiner *et al.* (*Proc. Natl. Acad. Sci. USA* 92:2755-2759, 1995) demonstrated that viral escape, in which the epitopes presented to PATR class I molecules mutated, was linked with a progression 25 toward chronic infection. These data show a role for the CTL in directing the course of HCV disease, and in shaping the genetic composition of HCV species in the persistently infected host.

In work in humans, Koziel and co-workers have established the presence of HCV-specific CTL in liver infiltrates from patients with chronic HCV infection (Koziel *et al.*, *J. Immunol.* 149:3339, 1992; and Koziel *et al.*, *J. Virol.* 67:7522, 1993), and have also identified a number of CTL epitopes recognized in the context of several different HLA 30 class I molecules. Other investigators have shown that HCV-specific CTL can be detected in the peripheral blood of patients with chronic hepatitis C (Cerny *et al.*, *J. Clin. Invest.* 95:521, 1995; Cerny *et al.*, *Curr. Topics in Micro. and Immunol.* 189:169, 1994; Cerny *et al.*, Abst. 2nd International Meeting on Hepatitis C and Related Viruses; La Jolla, CA, 1994; Battegay *et al.*, Abst. 2nd International Meeting on Hepatitis C and Related

Viruses; La Jolla, CA, 1994; Shirai *et al.*, *J. Virol.* 68:3334, 1994; Shirai *et al.*, *J. Immunol.* 154:2733, 1995; Battegay *et al.*, *J. Virol.* 69:2462, 1995). In addition, escape variants have been demonstrated in patients chronically infected with HCV (Chang *et al.*, *J. Clin. Invest.* 100:2376-2385, 1997; Tsai *et al.*, *Gastroenterology* 115:954-966, 1998).

5 The magnitude of the CTL responses observed in HCV-infected patients is, in general, higher than those observed in the case of chronic hepatitis B infection, suggesting that there is less impairment of specific T cell immunity than with HBV infection. The magnitude of CTL responses in HCV patients is, however, lower than those observed in HBV infected individuals who successfully cleared HBV infection.

10 10 These results support the understanding that HCV infected patients are capable of responding to active immunotherapy, and suggest that potentiation and increasing of T cell responses to HCV may be of use in therapy and prevention of chronic HCV infection (Prince, A. M. *FEMS Micro. Rev.* 14:273, 1994).

15 Several groups have analyzed the potential role of HCV-specific CTL responses in disease resistance and pathogenesis. In some studies no correlation was found between CTL viremia and CTL precursor frequency for individual HCV epitopes (Rehermann *et al.*, *J. Clin. Invest.* 98:1432-1440, 1996; Wong *et al.*, *J. Immunol.* 160:1479-1488, 1998). In other studies, however, it was shown that a clear correlation existed between levels of HCV infection and CTL responses, provided that the global response against multiple 20 CTL epitopes was considered (Rehermann *et al.*, *J. Virol.* 70:7092-7102, 1996). These data represent a strong rationale for development of vaccine constructs capable of inducing vigorous CTL responses directed against a multiplicity of conserved HCV-derived epitopes.

25 Koziel and colleagues have demonstrated the presence of HCV-specific CTLs, as well as T helper cell responses, in exposed but seronegative individuals (Koziel *et al.*, *J. Infect. Diseases* 176:859-866, 1997). In addition, HCV-specific CTLs have been detected in healthy, seronegative family members of chronically HCV-infected patients, indicating that a protective immunity is established in absence of a detectable infection (Bronowicki *et al.*, *J. Infect. Dis.* 176:518-522, 1997; Scognamiglio *et al.*, in preparation).

30 Experimental evidence also indicates that HTL epitopes play an important role in immune reactivity and defenses against HCV infection (Missale *et al.*, *J. Clin. Invest.* 98:706-714, 1996). Diepolder *et al.* (in *Lancet* 346:1006, 1995) have shown that a region of the NS3 gene (NS3 1007-1534) is recognized by patients who clear acute HCV infection, but is not seen by patients who develop chronic infection. Subsequent studies

showed that this particular region contain a highly cross-reactive HTL epitope (NS3 1248-1261), which binds with good affinity to 10 of 13 DR molecules tested, and is highly conserved in 30/33 different HCV isolates considered (Diepolder *et al.*, *J. Virol.* 71:6011-6019, 1997). These data suggested that directing HTL responses to this type of epitope (rather than to less cross-reactive and/or highly variable ones) will be of therapeutic and prophylactic benefit and strongly argue for inclusion of this and other epitopes with similar characteristics in HCV vaccine constructs.

The following examples illustrate identification, selection, and use of immunogenic Class I and Class II peptide epitopes for inclusion in vaccine compositions.

10

Example 1: HLA Class I and Class II Binding Assays

The following example of peptide binding to HLA molecules demonstrates quantification of binding affinities of HLA class I and class II peptides. Binding assays can be performed with peptides that are either motif-bearing or not motif-bearing.

15 Epstein-Barr virus (EBV)-transformed homozygous cell lines, fibroblasts, CIR, or 721.22 transfectants were used as sources of HLA class I molecules. The specific cell lines routinely used for purification of MHC class I and class II molecules are listed in Table XXIV. Cell lysates were prepared and HLA molecules purified in accordance with disclosed protocols (Sidney *et al.*, *Current Protocols in Immunology* 18.3.1 (1998);
20 Sidney, *et al.*, *J. Immunol.* 154:247 (1995); Sette, *et al.*, *Mol. Immunol.* 31:813 (1994)). HLA molecules were purified from lysates by affinity chromatography. The lysate was passed over a column of Sepharose CL-4B beads coupled to an appropriate antibody. The antibodies used for the extraction of HLA from cell lysates are listed in Table XXV. The anti-HLA column was then washed with 10mM Tris-HCL, pH 8.0, in 1% NP-40,
25 PBS, and PBS containing 0.4% n-octylglucoside and HLA molecules were eluted with 50mM diethylamine in 0.15M NaCl containing 0.4% n-octylglucoside, pH 11.5. A 1/25 volume of 2.0M Tris, pH 6.8, was added to the eluate to reduce the pH to ~8.0. Eluates were then be concentrated by centrifugation in Centriprep 30 concentrators (Amicon, Beverly, MA). Protein content was evaluated by a BCA protein assay (Pierce Chemical
30 Co., Rockford, IL) and confirmed by SDS-PAGE.

A detailed description of the protocol utilized to measure the binding of peptides to Class I and Class II MHC has been published (Sette *et al.*, *Mol. Immunol.* 31:813, 1994; Sidney *et al.*, in *Current Protocols in Immunology*, Margulies, Ed., John Wiley & Sons, New York, Section 18.3, 1998). Briefly, purified MHC molecules (5 to 500nM)

were incubated with various unlabeled peptide inhibitors and 1-10nM ^{125}I -radiolabeled probe peptides for 48h in PBS containing 0.05% Nonidet P-40 (NP40) (or 20% w/v digitonin for H-2 IA assays) in the presence of a protease inhibitor cocktail. All assays were at pH 7.0 with the exception of DRB1*0301, which was performed at pH 4.5, and

5 DRB1*1601 (DR2w21 β_1) and DRB4*0101 (DRw53), which were performed at pH 5.0.

Following incubation, MHC-peptide complexes were separated from free peptide by gel filtration on 7.8 mm x 15 cm TSK200 columns (TosoHaas 16215, Montgomeryville, PA). Because the large size of the radiolabeled peptide used for the DRB1*1501 (DR2w2 β_1) assay makes separation of bound from unbound peaks more difficult under these conditions, all DRB1*1501 (DR2w2 β_1) assays were performed using a 7.8mm x 30cm TSK2000 column eluted at 0.6 mls/min. The eluate from the TSK columns was passed through a Beckman 170 radioisotope detector, and radioactivity was plotted and integrated using a Hewlett-Packard 3396A integrator, and the fraction of peptide bound was determined.

15 Radiolabeled peptides were iodinated using the chloramine-T method.

Representative radiolabeled probe peptides utilized in each assay, and its assay specific IC₅₀ nM, are summarized in Tables IV and V. Typically, in preliminary experiments, each MHC preparation was titered in the presence of fixed amounts of radiolabeled peptides to determine the concentration of HLA molecules necessary to bind 10-20% of 20 the total radioactivity. All subsequent inhibition and direct binding assays were performed using these HLA concentrations.

Since under these conditions [label]<[HLA] and IC₅₀≥[HLA], the measured IC₅₀ values are reasonable approximations of the true K_D values. Peptide inhibitors are typically tested at concentrations ranging from 120 $\mu\text{g}/\text{ml}$ to 1.2 ng/ml, and are tested in 25 two to four completely independent experiments. To allow comparison of the data obtained in different experiments, a relative binding figure is calculated for each peptide by dividing the IC₅₀ of a positive control for inhibition by the IC₅₀ for each tested peptide (typically unlabeled versions of the radiolabeled probe peptide). For database purposes, and inter-experiment comparisons, relative binding values are compiled. These values 30 can subsequently be converted back into IC₅₀ nM values by dividing the IC₅₀ nM of the positive controls for inhibition by the relative binding of the peptide of interest. This method of data compilation has proven to be the most accurate and consistent for

comparing peptides that have been tested on different days, or with different lots of purified MHC.

Because the antibody used for HLA-DR purification (LB3.1) is α -chain specific, β_1 molecules are not separated from β_3 (and/or β_4 and β_5) molecules. The β_1 specificity of the binding assay is obvious in the cases of DRB1*0101 (DR1), DRB1*0802 (DR8w2), and DRB1*0803 (DR8w3), where no β_3 is expressed. It has also been demonstrated for DRB1*0301 (DR3) and DRB3*0101 (DR52a), DRB1*0401 (DR4w4), DRB1*0404 (DR4w14), DRB1*0405 (DR4w15), DRB1*1101 (DR5), DRB1*1201 (DR5w12), DRB1*1302 (DR6w19) and DRB1*0701 (DR7). The problem of β chain specificity for DRB1*1501 (DR2w2 β_1), DRB5*0101 (DR2w2 β_2), DRB1*1601 (DR2w21 β_1), DRB5*0201 (DR51Dw21), and DRB4*0101 (DRw53) assays is circumvented by the use of fibroblasts. Development and validation of assays with regard to DR β molecule specificity have been described previously (see, e.g., Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998).

Binding assays as outlined above may be used to analyze supermotif and/or motif-bearing epitopes as, for example, described in Example 2.

Example 2. Identification of Conserved HLA Supermotif- and Motif-Bearing CTL Candidate Epitopes

Vaccine compositions of the invention may include multiple epitopes that comprise multiple HLA supermotifs or motifs to achieve broad population coverage. This example illustrates the identification of supermotif- and motif-bearing epitopes for the inclusion in such a vaccine composition. Calculation of population coverage was performed using the strategy described below.

Computer searches and algorithms for identification of supermotif and/or motif-bearing epitopes

Computer searches for epitopes bearing HLA Class I or Class II supermotifs or motifs were performed as follows. All translated HCV isolate sequences were analyzed using a text string search software program, e.g., MotifSearch 1.4 (D. Brown, San Diego) to identify potential peptide sequences containing appropriate HLA binding motifs; alternative programs are readily produced in accordance with information in the art in view of the motif/supermotif disclosure herein. Furthermore, such calculations can be

made mentally. Identified A2-, A3-, and DR-supermotif sequences were scored using polynomial algorithms to predict their capacity to bind to specific HLA-Class I or Class II molecules. These polynomial algorithms take into account both extended and refined motifs (that is, to account for the impact of different amino acids at different positions), 5 and are essentially based on the premise that the overall affinity (or ΔG) of peptide-HLA molecule interactions can be approximated as a linear polynomial function of the type:

$$\text{"}\Delta G\text{"} = a_{1i} \times a_{2i} \times a_{3i} \dots \times a_{ni}$$

where a_{ji} is a coefficient which represents the effect of the presence of a given amino acid (j) at a given position (i) along the sequence of a peptide of n amino acids. The crucial 10 assumption of this method is that the effects at each position are essentially independent of each other (i.e., independent binding of individual side-chains). When residue j occurs at position i in the peptide, it is assumed to contribute a constant amount j_i to the free energy of binding of the peptide irrespective of the sequence of the rest of the peptide. This assumption is justified by studies from our laboratories that demonstrated that 15 peptides are bound to MHC and recognized by T cells in essentially an extended conformation (data omitted herein).

The method of derivation of specific algorithm coefficients has been described in Gulukota *et al.*, *J. Mol. Biol.* 267:1258-126, 1997; (see also Sidney *et al.*, *Human Immunol.* 45:79-93, 1996; and Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). 20 Briefly, for all i positions, anchor and non-anchor alike, the geometric mean of the average relative binding (ARB) of all peptides carrying j is calculated relative to the remainder of the group, and used as the estimate of j_i . For Class II peptides, if multiple alignments are possible, only the highest scoring alignment is utilized, following an iterative procedure. To calculate an algorithm score of a given peptide in a test set, the 25 ARB values corresponding to the sequence of the peptide are multiplied. If this product exceeds a chosen threshold, the peptide is predicted to bind. Appropriate thresholds are chosen as a function of the degree of stringency of prediction desired.

Selection of HLA-A2 supertype cross-reactive peptides

30 Complete polyprotein sequences from fourteen HCV isolates were aligned, then scanned, utilizing motif identification software, to identify conserved 9- and 10-mer sequences containing the HLA-A2-supermotif main anchor specificity.

A total of 231 conserved, HLA-A2 supermotif-positive sequences were identified. These peptides were then evaluated for the presence of A*0201 preferred secondary anchor residues using A*0201-specific polynomial algorithms. A total of 67 conserved, motif-bearing and algorithm-positive sequences were identified.

5 Fifty of these conserved, motif-containing 9- and 10-mer peptides were tested for their capacity to bind to purified HLA-A*0201 molecules *in vitro* (HLA-A*0201 is considered a prototype A2 supertype molecule). Sixteen peptides bound A*0201 with IC₅₀ values ≤500 nM; 4 with high binding affinities (IC₅₀ values ≤50 nM) and 12 with intermediate binding affinities, in the 50-500 nM range (Table XXVI).

10 These 16 peptides were then tested for binding to additional A2-supertype molecules (A*0202, A*0203, A*0206, and A*6802). As shown in Table XXVI, most of these peptides were found to be A2-supertype cross-reactive binders. More specifically, 12/16 (75%) peptides bound at least three of the five A2-supertype molecules tested.

15 *Selection of HLA-A3 supermotif-bearing epitopes*

The sequences from the same fourteen known HCV isolates scanned above were also examined for the presence of conserved peptides with the HLA-A3-supermotif primary anchors. A total of 71 conserved 9- or 10-mer motif containing sequences were identified. Further analysis using the A03 and A11 algorithms (see, e.g., Gulukota et al, 20 *J. Mol. Biol.* 267:1258-1267, 1997 and Sidney et al, *Human Immunol.* 45:79-93, 1996) identified 39 sequences that scored high in either or both algorithms. Twenty seven of the 39 peptides were synthesized and tested for binding to HLA-A*03 and HLA-A*11, the two most prevalent A3-supertype molecules. Fifteen peptides were identified which bound A3 and/or A11 with binding affinities of ≤500 nM (Table XXVII). These peptides 25 were then tested for binding cross-reactivity to the other common A3-supertype alleles (A*3101, A*3301, and A*6801). Seven of the 15 peptides bound at least three of the five HLA-A3-supertype molecules tested.

In the course of an independent series of experiments (Kubo et al., *J. Immunol.* 152:3913-3924, 1994), one peptide, HCV NS3 1262, not identified by the selection 30 criteria utilized above because it does not have the A3-supermotif main anchor specificity, was determined to be cross-reactive in the A3-supertype, binding A*03, A*11, and A*6801. It is also shown in Table XXVII. Interestingly, this peptide

represents a single residue N-terminal truncation of peptide 1073.14, which is also shown in Table XXVII.

In summary, 8 peptides that bind 3 or more A3-supertype molecules derived from conserved regions of the HCV genome were identified.

5

Selection of HLA-B7 supermotif bearing epitopes

When the same fourteen HCV isolates were also analyzed for the presence of conserved 9- or 10-mer peptides with the HLA-B7-supermotif, 35 sequences were identified. The corresponding peptides were synthesized and tested for binding to HLA-

10 B*0702, the most common B7-supertype allele (*i.e.*, the prototype B7 supertype allele). Thirteen peptides bound B*0702 with IC₅₀ of \leq 500 nM (Table XXVIIIa). These 13 peptides were then tested for binding to other common B7-supertype molecules (B*3501, B*51, B*5301, and B*5401). As shown in Table XXVIIIa, only 1 peptide (Core 169) was capable of binding to three or more of the five B7-supertype alleles tested.

15 To identify additional B7-supertype epitopes, further studies were undertaken. The protein sequences from the fourteen HCV isolates utilized above were again examined to identify conserved, motif-containing 8- and 11-mers. The isolates were also examined for 9- and 10-mer sequences allowing for lower conservancy (51%-78%). Twenty-five 8-mers, sixteen 11-mers, and thirty-five 9- and 10-mers were identified, 20 synthesized, and tested for binding to B*0702. Thirteen peptides bound with high or intermediate affinity (IC₅₀ \leq 500 nM) (Table XXVIIIb). These peptides were additionally screened for binding to other B7-supertype molecules. Only one cross-reactive binder, the NS3 1378 8-mer (peptide 29.0035/1260.04), was identified (Table XXVIIIb).

In summary, a total of two cross-reactive B7-supertype binders were identified 25 (Core 169 and NS3 1378).

Selection of A1 and A24 motif-bearing epitopes

To further increase population coverage, HLA-A1 and -A24 epitopes can also be incorporated into potential vaccine constructs.

30 In a previous analysis, two A1 and three A24 binders, 100% conserved among four strains of HCV, were identified (Wentworth *et al.*, *Int. Immunol.* 8:651-659, 1996). An analysis of the protein sequence data from the fourteen HCV strains utilized above demonstrated that these peptides were >79% conserved, and also identified an additional

eleven A1- and twenty five A24-motif-containing conserved sequences (see Table XXIXA and B). Eight of the additional eleven A1 peptides and seven of the additional twenty five A24 peptides were tested for binding to the appropriate HLA molecule (i.e., A1 or A24). Overall, as shown in Table XXIX, four A1-motif peptides (A) and three 5 A24-motif peptides (B) have been found with binding capacities of 500 nM or less for the appropriate allele-specific HLA molecule.

Analysis of the HLA-A2 and A3 supermotif-bearing epitopes identified above revealed that in 13/14 cases, peptides binding the supertype prototype HLA molecule (i.e. A*0201 for the A2 supertype, and A*0301 for the A3 supertype) with an IC₅₀ of less than 10 100nM were cross-reactive and recognized by HCV-infected patients as described in Example 3, which follows. Based on these observations, two A1 peptides and one A24 peptide epitopes were also selected as candidates for inclusion in vaccine compositions; these peptides bind the appropriate HLA molecule with an IC₅₀ of less than 100nM.

15 **Example 3: Confirmation of Immunogenicity**

*Evaluation of A*0201 immunogenicity*

It has been shown that CTL induced in A*0201/K^b transgenic mice exhibit specificity similar to CTL induced in the human system (see, e.g., Vitiello *et al.*, *J. Exp. Med.* 173:1007-1015, 1991; Wentworth *et al.*, *Eur. J. Immunol.* 26:97-101, 1996). 20 Accordingly, these mice were used to evaluate the immunogenicity of the twelve conserved A2-supertype cross-reactive peptides identified in Example 2 above.

CTL induction in transgenic mice following peptide immunization has been described (Vitiello *et al.*, *J. Exp. Med.* 173:1007-1015, 1991; Alexander *et al.*; *J. Immunol.* 159:4753-4761, 1997). In these studies, mice were injected subcutaneously at 25 the base of the tail with each peptide (50 µg/mouse) emulsified in IFA in the presence of an excess of an IA^b-restricted helper peptide (140 µg/mouse) (HBV core 128-140, Sette *et al.*, *J. Immunol.* 153:5586-5592, 1994). Eleven days after injection, splenocytes were incubated in the presence of peptide-loaded syngenic LPS blasts. After six days, cultures were assayed for cytotoxic activity using peptide-pulsed targets. The data, summarized in 30 Table XXX, indicate that 7 of the 12 peptides (58%) were capable of inducing primary CTL responses in A*0201/K^b transgenic mice. (For these studies, a peptide was considered positive if it induced CTL (L.U. 30/10⁶ cells ≥2 in at least two transgenic animals (Wentworth *et al.*, *Eur. J. Immunol.* 26:97-101, 1996).

The conserved, cross reactive candidate CTL epitopes were also tested for recognition *in vitro* by PBMCs obtained from HCV-infected patients. Briefly, PBMC from patients infected with HCV were cultured in the presence of 10 µg/ml of synthetic peptide. After 7 and 14 days, the cultures were restimulated with peptide. The cultures 5 were assayed for cytolytic activity on day 21 using target cells pulsed with the specific peptide in a standard four hour ⁵¹Cr release assay. The data are summarized in Table XXX. As shown, all 12 peptides are CTL epitopes recognized by PBMC from HCV-infected patients. From the data in Table XXX, it is interesting to note that HLA 10 transgenics did not fully reveal the immunogenicity of some peptides that were positive in recall responses. This apparent discrepancy may reflect differences in the route of immunization utilized (e.g., natural infection versus peptide immunization), or CTL repertoire.

*Evaluation of A*03/A11 immunogenicity*

15 The immunogenicity of six of the eight A3-supertype cross-reactive peptides identified in Example 2 above was evaluated in HLA-A11/K^b transgenic mice, using the protocol described above for HLA-A2 transgenic mice (Alexander *et al.*, *J. Immunol.* 159:4753-4761, 1997). Five of these six peptides were able to induce primary CTL responses (Table XXXI).

20 All eight peptides were also studied by collaborators using PBMC cultures from HCV infected patients and contacts of such patients. This data is also summarized in Table XXXI. Briefly, all eight peptides were recognized by HCV infected individuals.

Evaluation of B7 immunogenicity

25 One of the two B7-supertype cross-reactive peptides (1145.12, Core 169) has been evaluated for immunogenicity in HCV-infected patients. Two independent collaborators have shown that this peptide is indeed immunogenic, and is recognized by T cells from HCV-infected patients (Chang *et al.*, *J. Immunol.* 162:1156-1164, 1999)

30 **Example 4: Implementation of the Extended Supermotif to Improve the Binding Capacity of Native Epitopes by Creating Analogs**

HLA motifs and supermotifs (comprising primary and/or secondary residues) are useful in the identification and preparation of highly cross-reactive native peptides, as demonstrated herein. Moreover, the definition of HLA motifs and supermotifs also

allows one to engineer highly cross-reactive epitopes by identifying residues within a native peptide sequence which can be analogued, or "fixed" to confer upon the peptide certain characteristics, e.g. greater cross-reactivity within the group of HLA molecules that comprise a supertype, and/or greater binding affinity for some or all of those HLA 5 molecules. Examples of analog peptides that exhibit modulated binding affinity are set forth in this example.

Analoging at Primary Anchor Residues

As shown in Example 2, more than ten different HCV-derived, A2-supertype-restricted epitopes were identified. Peptide engineering strategies are implemented to further increase the cross-reactivity of the candidate epitopes identified above which bind 3/5 of the A2 supertype alleles tested. On the basis of the data disclosed, e.g., in related and co-pending U.S.S.N 09/226,775, the main anchors of A2-supermotif-bearing peptides are altered, for example, to introduce a preferred L, I, V, or M at position 2, and I or V at 15 the C-terminus.

To analyze the cross-reactivity of the analog peptides, each engineered analog is initially tested for binding to the prototype A2 supertype allele A*0201, then, if A*0201 binding capacity is maintained, for A2-supertype cross-reactivity.

Similarly, analogs of HLA-A3 supermotif-bearing epitopes may also be 20 generated. For example, peptides binding to 3/5 of the A3-supertype molecules may be engineered at primary anchor residues to possess a preferred residue (V, S, M, or A) at position 2.

The analog peptides are then tested for the ability to bind A*03 and A*11 (prototype A3 supertype alleles). Those peptides that demonstrate ≤ 500 nM binding 25 capacity are then tested for A3-supertype cross-reactivity.

Similarly to the A2- and A3- motif bearing peptides, peptides binding 3 or more B7-supertype alleles may be improved, where possible, to achieve increased cross-reactive binding. B7 supermotif-bearing peptides may, for example, be engineered to possess a preferred residue (V, I, L, or F) at the C-terminal primary anchor position, as 30 demonstrated by Sidney *et al.* (*J. Immunol.* 157:3480-3490, 1996).

Analoging at Secondary Anchor Residues

Moreover, HLA supermotifs are of value in engineering highly cross-reactive peptides and/or peptides that bind HLA molecules with increased affinity by identifying

particular residues at secondary anchor positions that are associated with such properties. Demonstrating this, the binding capacity of a peptide representing a discreet single amino acid substitution at position one was analyzed. Peptide 1145.13 (Table XXVIIc), which represents the substitution of L to F at position 1 of the core 169 sequence, binds all five 5 B7-supertype molecules with a good affinity (all IC_{50} values ≤ 132 nM), and in 3 instances has higher affinity over that of the parent peptide by >35-fold.

Because so few B7-supertype cross-reactive epitopes were identified, our results from previous binding evaluations were analyzed to identify conserved (8-, 9-, 10-, or 11-mer) peptides which bind, minimally, 3/5 B7 supertype molecules with weak affinity 10 (IC_{50} of 500nM-5 μ M). This analysis identified 9 peptides, 6 of which are analogued (including core 169 which had been previously analogued). These peptides are tested for enhanced binding affinity and B7-supertype cross-reactivity.

Engineered analogs with sufficiently improved binding capacity or cross-reactivity are tested for immunogenicity in HLA-B7-transgenic mice, following for 15 example, IFA immunization or lipopeptide immunization.

In conclusion, these data demonstrate that by the use of even single amino acid substitutions, it is possible to increase the binding affinity and/or cross-reactivity of peptide ligands for HLA supertype molecules.

20 Example 5: Identification of conserved HCV-derived sequences with HLA-DR binding motifs

Peptide epitopes bearing an HLA class II supermotif or motif may also be identified as outlined below using methodology similar to that described in Examples 1-3.

25 *Selection of HLA-DR-supermotif-bearing epitopes*

To identify HCV-derived, HLA class II HTL epitopes, the same fourteen HCV polyprotein sequences used for the identification of HLA Class I supermotif/motif sequences were analyzed for the presence of sequences bearing an HLA-DR-motif or supermotif. Specifically, 15-mer sequences were selected comprising a DR-supermotif, 30 further comprising a 9-mer core, and three-residue N- and C-terminal flanking regions (15 amino acids total). It was also required that the 15-mer sequence be conserved in at least 79% (11/14) of the HCV strains analyzed. These criteria identified a total of 49 non-redundant sequences, which are shown in Table XXXIIA. (In the context of Class II

epitopes, a sequence is considered operationally redundant if more than 80% of it's sequence overlaps with another peptide.)

Protocols for predicting peptide binding to DR molecules have been developed (Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). These protocols, specific for 5 individual DR molecules, allow the scoring, and ranking, of 9-mer core regions. Each protocol not only scores peptide sequences for the presence of DR-supermotif primary anchors (i.e., at position 1 and position 6) within a 9-mer core, but additionally evaluates sequences for the presence of secondary anchors. Using allele specific selection tables (see, *e.g.*, Southwood *et al.*, *ibid.*), it has been found that these protocols efficiently select 10 peptide sequences with a high probability of binding a particular DR molecule.

Additionally, it has been found that performing these protocols in tandem, specifically those for DR1, DR4w4, and DR7, can efficiently select DR cross-reactive peptides.

To see if these protocols serve to identify additional epitopes, the same HCV polyproteins used above were re-scanned for the presence of 15-mer peptides with 9-mer 15 core regions that were $\geq 79\%$ (11/14 strains) conserved. This identified 152 sequences; 49 of which were identified previously, as described above. Next, the 9-mer core region of each of these peptides was scored using the DR1, DR4w4, and DR7 algorithms. Twenty-two peptides, including 12 new sequences (10 peptides were from the original set of 49) were found to have 9-mer cores with protocol-derived scores predictive of cross-reactive 20 DR binders. The 12 additional sequences are shown in Table XXXIIB.

The conserved, HCV-derived peptides identified above were tested for their binding capacity for various common HLA-DR molecules. All peptides were initially tested for binding to the DR molecules in the primary panel: DR1, DR4w4, and DR7. Peptides binding at least 2 of these 3 DR molecules were then tested for binding to 25 DR2w2 β 1, DR2w2 β 2, DR6w19, and DR9 molecules in secondary assays. Finally, peptides binding at least 2 of the 4 secondary panel DR molecules, and thus cumulatively at least 4 of 7 different DR molecules, were screened for binding to DR4w15, DR5w11, and DR8w2 molecules in tertiary assays. Peptides binding at least 7 of the 10 DR molecules comprising the primary, secondary, and tertiary screening assays were 30 considered cross-reactive DR binders. The composition of these screening panels, and the phenotypic frequency of associated antigens, are shown in Table XXXIII.

Upon testing, it was found that 29 of the original 75 peptides (39%) bound two or more of the primary HLA molecules. Twenty-six of these cross-reactive binders were

then tested in the secondary assays, and nineteen were found to bind at least four of the seven HLA DR molecules in the primary and secondary panels. Finally, the nineteen peptides passing the secondary screening phase were tested for binding in the tertiary assays. As a result, nine peptides were identified which bound at least seven of ten common HLA-DR molecules. Table XXXIV shows these nine peptides and their binding capacity for each allele-specific HLA-DR molecule in the primary through tertiary panels. Also shown in Table XXXIV are two peptides (F134.05 and F134.08) for which a complete binding analysis was not performed. However, both of these peptides bound six of the seven HLA DR molecules tested. F134.08 nests peptide 1283.44, which bound eight of 10 allele-specific HLA molecules.

In conclusion, eleven cross-reactive DR-binding peptides, derived from six discrete (*i.e.* non-redundant) regions of the HCV genome, have been identified. Two of the six regions from which these epitopes were derived are covered by multiple, overlapping epitopes.

15

Selection of conserved DR3 motif peptides

Because HLA-DR3 is an allele that is prevalent in Caucasian, Black, and Hispanic populations, DR3 binding capacity is an important criterion in the selection of HTL epitopes. However, data generated previously indicated that DR3 only rarely cross-reacts with other DR alleles (Sidney *et al.*, *J. Immunol.* 149:2634-2640, 1992; Geluk *et al.*, *J. Immunol.* 152:5742-5748, 1994; Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). This is not entirely surprising in that the DR3 peptide-binding motif appears to be distinct from the specificity of most other DR alleles.

To efficiently identify peptides that bind DR3, target proteins were analyzed for conserved sequences carrying one of the two DR3 specific binding motifs reported by Geluk *et al.* (*J. Immunol.* 152:5742-5748, 1994). Fifteen sequences, including a peptide nested within a DR-supermotif sequence identified above (peptide Pape 22), were identified (Table XXXIId). Preferably, DR3 motifs will be found clustered in proximity with DR supermotif regions.

30 Fourteen of the fifteen peptides containing a DR3 motif were tested for their DR3 binding capacity. Two peptides (CH35.0106 and CH35.0107) were found to bind DR3 with an affinity of 1 μ M or less (Table XXXV), and thereby qualify as HLA class II high affinity binders.

DR3 binding epitopes identified in this manner may then be included in vaccine compositions with DR supermotif-bearing peptide epitopes.

5 **Example 6: Immunogenicity of candidate HCV-derived HTL epitopes and known dominant HCV HTL epitope**

In the course of collaborative studies with G. Pape and C. Ferrari, eight conserved, HCV-derived peptides have been identified which are recognized by HCV-infected individuals.

One of these studies (Diepolder *et al.*, *J. Virol.* 71:6011-6019, 1997), identified 10 peptide F98.05, which spans residues 1248-1261 of the NS3 protein, as an immunodominant CD4+ T-cell epitope that was recognized by 14/23 NS3-specific CD4+ T-cell clones from 4/5 patients with acute hepatitis C infection. This epitope, shown above to be an HLA-DR cross-reactive binder (see Table XXXIV), was capable of being presented to helper CD4+ T cells by multiple HLA molecules (DR4, DR11, DR12, DR13, 15 and DR16). Two other peptides, Pape 22 and Pape 29, were also recognized by CD4+ T cell clones, although, in a more limited context; correspondingly, neither of these peptides are DR-cross-reactive binders.

By direct peripheral blood T cell stimulation and by fine specificity analysis of HCV-specific T-cell lines and clones, studies done in collaboration with Ferrari's group 20 identified 6 immunodominant epitopes, including one also identified in the Pape collaboration, that are derived from conserved regions of the core, NS3, and NS4 proteins. These epitopes were also found to be cross-reactive, being presented to T cells in the context of different Class II molecules. Three of the 6 epitopes, F98.04 (F134.03), F134.05 and F134.08, are cross-reactive HLA-DR binders (see Table XXXIV).

25 In conclusion, the immunogenicity of 8 epitopes derived from conserved regions of the HCV genome has been demonstrated. Three of these epitopes (F98.05, F134.05, and F134.08; see Table XXXIV) are broadly cross-reactive HLA-DR binding peptides.

30 **Example 7. Calculation of phenotypic frequencies of HLA-supertypes in various ethnic backgrounds to determine breadth of population coverage**

This example illustrates the assessment of the breadth of population coverage of a vaccine composition comprised of multiple epitopes comprising multiple supermotifs and/or motifs.

In order to analyze population coverage, gene frequencies of HLA alleles were determined. Gene frequencies for each HLA allele were calculated from antigen or allele frequencies utilizing the binomial distribution formulae $gf=1-(\text{SQRT}(1-af))$ (see, e.g., Sidney *et al.*, *Human Immunol.* 45:79-93, 1996). To obtain overall phenotypic frequencies, cumulative gene frequencies were calculated, and the cumulative antigen frequencies derived by the use of the inverse formula $[af=1-(1-Cgf)^2]$.

Where frequency data was not available at the level of DNA typing, correspondence to the serologically defined antigen frequencies was assumed. To obtain total potential supertype population coverage no linkage disequilibrium was assumed, and only alleles confirmed to belong to each of the supertypes were included (minimal estimates). Estimates of total potential coverage achieved by inter-loci combinations were made by adding to the A coverage the proportion of the non-A covered population that could be expected to be covered by the B alleles considered (e.g., $\text{total}=A+B^*(1-A)$). Confirmed members of the A3-like supertype are A3, A11, A31, A*3301, and A*6801. Although the A3-like supertype may also include A34, A66, and A*7401, these alleles were not included in overall frequency calculations. Likewise, confirmed members of the A2-like supertype family are A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*6802, and A*6901. Finally, the B7-like supertype-confirmed alleles are: B7, B*3501-03, B51, B*5301, B*5401, B*5501-2, B*5601, B*6701, and B*7801 (potentially also B*1401, B*3504-06, B*4201, and B*5602).

Population coverage achieved by combining the A2-, A3- and B7-supertypes is approximately 86% in five major ethnic groups (see Table XXI). Coverage may be extended by including peptides bearing the A1 and A24 motifs. On average, A1 is present in 12% and A24 in 29% of the population across five different major ethnic groups (Caucasian, North American Black, Chinese, Japanese, and Hispanic). Together, these alleles are represented with an average frequency of 39% in these same ethnic populations. The total coverage across the major ethnicities when A1 and A24 are combined with the coverage of the A2-, A3- and B7-supertype alleles is >95%. An analogous approach can be used to estimate population coverage achieved with combinations of class II motif-bearing epitopes.

Summary of candidate HLA class I and class II epitopes

In summary, on the basis of the data presented in the above examples, 26 CTL candidate peptide epitopes derived from conserved regions of the HCV virus have been

identified (Table XXXVIa). These include twelve HLA-A2 supermotif-bearing epitopes, eight HLA-A3 supermotif-bearing epitopes, and one HLA-B7 supermotif-bearing epitope, each capable of binding to multiple A2-, A3-, or B7-supertype molecules, and immunogenic in HLA transgenic mice or antigenic for human PBL (with the exception of 5 peptide 29.0035/1260.04). Additional epitopes not evaluated for immunogenicity are also included. They are an additional B7-supermotif-bearing epitope and two HLA-A1 and one HLA-A24 high-affinity binding peptides. A known HLA-A31 restricted epitope (VGIYLLPNR), which also binds HLA-A33, is also set out in Table XXXVIa and is useful in combination with other Class I or Class II epitopes.

10 With these 26 CTL epitopes (as disclosed herein and from the art), average population coverage, (*i.e.*, recognition of at least one HCV epitope), is predicted to be greater than 95% in each of five major ethnic populations. The potential redundancy of coverage afforded by 25 of these epitopes (the peptide 24.0086 was not included) was estimated using the game theory Monte Carlo simulation analysis, which is known in the 15 art (see *e.g.*, Osborne, M.J. and Rubinstein, A. "A course in game theory" MIT Press, 1994). As shown in Figure 1, it is estimated that 90% of the individuals in a population comprised of the Caucasian, North American Black, Japanese, Chinese, and Hispanic ethnic groups would recognize 2 or more of the candidate epitopes described herein.

10 A list of HCV-derived HTL epitopes that would be preferred for use in the design 20 of minigene constructs or other vaccine formulations is summarized in Table XXXVIb. As shown, 9 different peptide-binding regions have been identified which bind multiple HLA-DR molecules or bind HLA-DR3. (In the case of the NS4 1914-1935 region, the longer peptide, F134.08, recognized by patients, was chosen over the shorter peptide, 1283.44. The longer peptide essentially incorporates the shorter peptide, and also binds 25 additional DR molecules that the shorter peptide does not bind.) Three of these peptides have been recognized as dominant epitopes in HCV infected patients.

10 It is estimated that each of 10 common DR molecules recognizing the DR 20 supermotif, and DR3, are covered by a minimum of 2 epitopes. Correspondingly, the total estimated population coverage represented by this panel of epitopes is in excess of 25 30 91% in each of the 5 major ethnic populations (Table XXXVII).

Example 8: Recognition Of Generation Of Endogenous Processed Antigens After Priming

This example determines that CTL induced by native or analogued peptide epitopes identified and selected as described in Examples 1-6 recognize endogenously synthesized, *i.e.*, native antigens.

Effector cells isolated from transgenic mice that are immunized with peptide epitopes as in Example 3, for example HLA-A2 supermotif-bearing epitopes, are re-stimulated *in vitro* using peptide-coated stimulator cells. Six days later, effector cells are assayed for cytotoxicity and the cell lines that contain peptide-specific cytotoxic activity 10 are further re-stimulated. An additional six days later, these cell lines are tested for cytotoxic activity on ⁵¹Cr labeled Jurkat-A2.1/K^b target cells in the absence or presence of peptide, and also tested on ⁵¹Cr labeled target cells bearing the endogenously synthesized antigen, *i.e.* cells that are stably transfected with HCV expression vectors.

The result will demonstrate that CTL lines obtained from animals primed with 15 peptide epitope recognize endogenously synthesized HCV antigen. The choice of transgenic mouse model to be used for such an analysis depends upon the epitope(s) that is being evaluated. In addition to HLA-A*0201/K^b transgenic mice, several other transgenic mouse models including mice with human A11, which may also be used to evaluate A3 epitopes, and B7 alleles have been characterized and others (*e.g.*, transgenic 20 mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed, which may be used to evaluate HTL epitopes.

Example 9: Activity Of CTL-HTL Conjugated Epitopes In Transgenic Mice

This example illustrates the induction of CTLs and HTLs in transgenic mice by 25 use of an HCV CTL/HTL peptide conjugate whereby the vaccine composition comprises peptides administered to an HCV-infected patient or an individual at risk for HCV. The peptide composition can comprise multiple CTL and/or HTL epitopes. This analysis demonstrates enhanced immunogenicity that can be achieved by inclusion of one or more HTL epitopes in a vaccine composition. Such a peptide composition can comprise a 30 lipidated HTL epitope conjugated to a preferred CTL epitope containing, for example, at least one CTL epitope selected from Table XXVI-XXIX, or an analog of that epitope. The HTL epitope is, for example, selected from Table XXXII.

Lipopeptide preparation: Lipopeptides are prepared by coupling the appropriate fatty acid to the amino terminus of the resin bound peptide. A typical procedure is as

follows: A dichloromethane solution of a four-fold excess of a pre-formed symmetrical anhydride of the appropriate fatty acid is added to the resin and the mixture is allowed to react for two hours. The resin is washed with dichloromethane and dried. The resin is then treated with trifluoroacetic acid in the presence of appropriate scavengers [e.g. 5% 5 (v/v) water] for 60 minutes at 20°C. After evaporation of excess trifluoroacetic acid, the crude peptide is washed with diethyl ether, dissolved in methanol and precipitated by the addition of water. The peptide is collected by filtration and dried.

10 Immunization procedures: Immunization of transgenic mice is performed as described (Alexander *et al.*, *J. Immunol.* 159:4753-4761, 1997). For example, A2/K^b mice, which are transgenic for the human HLA A2.1 allele and are useful for the assessment of the immunogenicity of HLA-A*0201 motif- or HLA-A2 supermotif-bearing epitopes, are primed subcutaneously (base of the tail) with 0.1 ml of peptide conjugate formulated in saline, or DMSO/saline. Seven days after priming, splenocytes obtained from these animals are restimulated with syngenic irradiated LPS-activated 15 lymphoblasts coated with peptide.

Cell lines: Target cells for peptide-specific cytotoxicity assays are Jurkat cells transfected with the HLA-A2.1/K^b chimeric gene (e.g., Vitiello *et al.*, *J. Exp. Med.* 173:1007, 1991)

20 *In vitro* CTL activation: One week after priming, spleen cells (30x10⁶ cells/flask) are co-cultured at 37°C with syngeneic, irradiated (3000 rads), peptide coated lymphoblasts (10x10⁶ cells/flask) in 10 ml of culture medium/T25 flask. After six days, effector cells are harvested and assayed for cytotoxic activity.

25 Assay for cytotoxic activity: Target cells (1.0 to 1.5x10⁶) are incubated at 37°C in the presence of 200 µl of ⁵¹Cr. After 60 minutes, cells are washed three times and resuspended in R10 medium. Peptide is added where required at a concentration of 1 µg/ml. For the assay, 10⁴ ⁵¹Cr-labeled target cells are added to different concentrations of effector cells (final volume of 200 µl) in U-bottom 96-well plates. After a 6 hour incubation period at 37°C, a 0.1 ml aliquot of supernatant is removed from each well and radioactivity is determined in a Micromedic automatic gamma counter. The percent 30 specific lysis is determined by the formula: percent specific release = 100 x (experimental release - spontaneous release)/(maximum release - spontaneous release). To facilitate comparison between separate CTL assays run under the same conditions, % ⁵¹Cr release data is expressed as lytic units/10⁶ cells. One lytic unit is arbitrarily defined as the number of effector cells required to achieve 30% lysis of 10,000 target cells in a 6

hour ^{51}Cr release assay. To obtain specific lytic units/ 10^6 , the lytic units/ 10^6 obtained in the absence of peptide is subtracted from the lytic units/ 10^6 obtained in the presence of peptide. For example, if 30% ^{51}Cr release is obtained at the effector (E): target (T) ratio of 50:1 (i.e., 5×10^5 effector cells for 10,000 targets) in the absence of peptide and 5:1 (i.e., 5 $\times 10^4$ effector cells for 10,000 targets) in the presence of peptide, the specific lytic units would be: $[(1/50,000)-(1/500,000)] \times 10^6 = 18 \text{ LU}$.

The results are analyzed to assess the magnitude of the CTL responses of animals injected with the immunogenic CTL/HTL conjugate vaccine preparation and are compared to the magnitude of the CTL response achieved using the CTL epitope as outlined in Example 3. Analyses similar to this may be performed to evaluate the immunogenicity of peptide conjugates containing multiple CTL epitopes and/or multiple HTL epitopes. In accordance with these procedures it is found that a CTL response is induced, and concomitantly that an HTL response is induced upon administration of such compositions.

15

Example 10. Selection of CTL and HTL epitopes for inclusion in an HCV-specific vaccine.

This example illustrates the procedure for the selection of peptide epitopes for vaccine compositions of the invention. The peptides in the composition can be in the form of a nucleic acid sequence, either single or one or more sequences (i.e., minigene) that encodes peptide(s), or may be single and/or polyepitopic peptides.

Epitopes are selected which, upon administration, mimic immune responses that have been observed to be correlated with tumor clearance. For example, vaccine can include 3-4 epitopes that come from at least one HCV antigen region. Epitopes from one region can be used in combination with epitopes from one or more additional HCV antigen regions. Analogs of epitopes can also be selected for inclusion in the vaccine.

Epitopes are often selected that have a binding affinity of an IC_{50} of 500 nM or less for an HLA class I molecule, or for class II, an IC_{50} of 1000 nM or less.

Sufficient supermotif bearing peptides, or a sufficient array of allele-specific motif bearing peptides, are selected to give broad population coverage. For example, epitopes are selected to provide at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess breadth, or redundancy, of population coverage.

When creating a polyepitopic compositions, *e.g.* a minigene, it is typically desirable to generate the smallest peptide possible that encompasses the epitopes of interest. The principles employed are similar, if not the same, as those employed when selecting a peptide comprising nested epitopes. Additionally, however, upon 5 determination of the nucleic acid sequence to be provided as a minigene, the peptide sequence encoded thereby is analyzed to determine whether any "junctional epitopes" have been created. A junctional epitope is a potential HLA binding epitope, as predicted, *e.g.*, by motif analysis. Junctional epitopes are generally to be avoided because the recipient may bind to an HLA molecule and generate an immune response to that epitope, 10 which is not present in a native protein sequence.

Peptide epitopes for inclusion in vaccine compositions are, for example, selected from those listed in Tables XXVI-XXIX and Table XXXII. A vaccine composition comprised of selected peptides, when administered, is safe, efficacious, and elicits an immune response similar in magnitude of an immune response that clears an acute HCV 15 infection.

Example 11: Construction of Minigene Multi-Epitope DNA Plasmids

This example provides guidance for the construction of a minigene expression plasmid. Minigene plasmids may, of course, contain various configurations of CTL 20 and/or HTL epitopes or epitope analogs as described herein. Examples of the construction and evaluation of expression plasmids are described, for example, in co-pending U.S.S.N. 09/311,784 filed 5/13/99. An example of such a plasmid for the expression of HCV epitopes is shown in Figure 2, which illustrates the orientation of HCV peptide epitopes in a minigene construct.

25 A minigene expression plasmid may include multiple CTL and HTL peptide epitopes. In the present example, HLA-A2, -A3, -B7 supermotif-bearing peptide epitopes and HLA-A1 and -A24 motif-bearing peptide epitopes are used in conjunction with DR supermotif-bearing epitopes and/or DR3 epitopes (Figure 2). Preferred epitopes are identified, for example, in Tables XXVI-XXIX and XXXII. HLA class I supermotif or 30 motif-bearing peptide epitopes derived from multiple HCV antigens, *e.g.*, the core, NS4, NS3, NS5, NS1/E2, are selected such that multiple supermotifs/motifs are represented to ensure broad population coverage. Similarly, HLA class II epitopes are selected from multiple HCV antigens to provide broad population coverage, *i.e.* both HLA DR-1-4-7 supermotif-bearing epitopes and HLA DR-3 motif-bearing epitopes are selected for

inclusion in the minigene construct. The selected CTL and HTL epitopes are then incorporated into a minigene for expression in an expression vector.

This example illustrates the methods to be used for construction of such a minigene-bearing expression plasmid. Other expression vectors that may be used for 5 minigene compositions are available and known to those of skill in the art.

The minigene DNA plasmid contains a consensus Kozak sequence and a consensus murine kappa Ig-light chain signal sequence followed by CTL and/or HTL epitopes selected in accordance with principles disclosed herein. The sequence encodes an open reading frame fused to the Myc and His antibody epitope tag coded for by the 10 pcDNA 3.1 Myc-His vector.

Overlapping oligonucleotides, for example eight oligonucleotides, averaging approximately 70 nucleotides in length with 15 nucleotide overlaps, are synthesized and HPLC-purified. The oligonucleotides encode the selected peptide epitopes as well as appropriate linker nucleotides, Kozak sequence, and signal sequence. The final 15 multiepitope minigene is assembled by extending the overlapping oligonucleotides in three sets of reactions using PCR. A Perkin/Elmer 9600 PCR machine is used and a total of 30 cycles are performed using the following conditions: 95°C for 15 sec, annealing temperature (5° below the lowest calculated Tm of each primer pair) for 30 sec, and 72°C for 1 min.

20 For the first PCR reaction, 5 µg of each of two oligonucleotides, *i.e.*, an amplification primer pair, are annealed and extended: Oligonucleotides 1+2, 3+4, 5+6, and 7+8 are combined in 100 µl reactions containing *Pfu* polymerase buffer (1x= 10 mM KCL, 10 mM (NH₄)₂SO₄, 20 mM Tris-chloride, pH 8.75, 2 mM MgSO₄, 0.1% Triton X-100, 100 µg/ml BSA), 0.25 mM each dNTP, and 2.5 U of *Pfu* polymerase. The full-length dimer products are gel-purified, and two reactions containing the product of 1+2 and 3+4, and the product of 5+6 and 7+8 are mixed, annealed, and extended for 10 cycles. Half of the two reactions are then mixed, and 5 cycles of annealing and extension carried out before flanking primers are added to amplify the full length product for 25 additional cycles. The full-length product is gel-purified and cloned into pCR-blunt 25 (Invitrogen) and individual clones are screened by sequencing.

Example 12. The plasmid construct and the degree to which it induces immunogenicity.

The degree to which the plasmid construct prepared using the methodology outlined in Example 11 is able to induce immunogenicity is evaluated through *in vivo*

injections into mice and subsequent *in vitro* assessment of CTL and HTL activity, which are analysed using cytotoxicity and proliferation assays, respectively, as detailed *e.g.*, in U.S.S.N. 09/311,784 filed 5/13/99 and Alexander *et al.*, *Immunity* 1:751-761, 1994. For example, to assess the capacity of a pMin minigene construct that contains HLA-A2 supermotif epitopes to induce CTLs *in vivo*, HLA-A2.1/K^b transgenic mice are immunized intramuscularly with 100 µg of naked cDNA. As a means of comparing the level of CTLs induced by cDNA immunization, a control group of animals is also immunized with an actual peptide composition that comprises multiple epitopes synthesized as a single polypeptide as they would be encoded by the minigene.

10 Splenocytes from immunized animals are stimulated twice with each of the respective compositions (peptide epitopes encoded in the minigene or the polyepitopic peptide), then assayed for peptide-specific cytotoxic activity in a ⁵¹Cr release assay. The results indicate the magnitude of the CTL response directed against the A3-restricted epitope, thus indicating the *in vivo* immunogenicity of the minigene vaccine and

15 polyepitopic vaccine. It is, therefore, found that the minigene elicits immune responses directed toward the HLA-A2 supermotif peptide epitopes as does the polyepitopic peptide vaccine. A similar analysis is also performed using other HLA-A3 and HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 and HLA-B7 motif or supermotif epitopes.

20 To assess the capacity of a class II epitope encoding minigene to induce HTLs *in vivo*, I-A^b restricted mice, for example, are immunized intramuscularly with 100 µg of plasmid DNA. As a means of comparing the level of HTLs induced by DNA immunization, a group of control animals is also immunized with an actual peptide composition emulsified in complete Freund's adjuvant.

25 CD4+ T cells, *i.e.* HTLs, are purified from splenocytes of immunized animals and stimulated with each of the respective compositions (peptides encoded in the minigene). The HTL response is measured using a ³H-thymidine incorporation proliferation assay, (*see, e.g.*, Alexander *et al.* *Immunity* 1:751-761, 1994). the results indicate the magnitude of the HTL response, thus demonstrating the *in vivo* immunogenicity of the minigene.

30 Alternatively, plasmid constructs can be evaluated *in vitro* by testing for epitope presentation by APC following transduction or transfection of the APC with an epitope-expressing nucleic acid construct. Such a study determines "antigenicity" and allows the use of human APC. The assay determines the ability of the epitope to be presented by the

APC in a context that is recognized by a T cell by quantifying the density of epitope-HLA class I complexes on the cell surface. Quantitation can be performed by directly measuring the amount of peptide eluted from the APC (see, e.g., Sijts *et al.*, *J. Immunol.* 156:683-692, 1996; Demotz *et al.*, *Nature* 342:682-684, 1989); or the number of peptide-
5 HLA class I complexes can be estimated by measuring the amount of lysis or lymphokine release induced by infected or transfected target cells, and then determining the concentration of peptide necessary to obtain equivalent levels of lysis or lymphokine release (see, e.g., Kageyama *et al.*, *J. Immunol.* 154:567-576, 1995).

10 Example 13: Peptide Composition for Prophylactic Uses

Vaccine compositions of the present invention are used to prevent HCV infection in persons who are at risk for such infection. For example, a polyepitopic peptide epitope composition (or a nucleic acid comprising the same) containing multiple CTL and HTL epitopes such as those selected in Examples 9 and/or 10, which are also selected to target
15 greater than 80% of the population, is administered to individuals at risk for HCV infection. The composition is provided as a single lipidated polypeptide that encompasses multiple epitopes. The vaccine is administered in an aqueous carrier comprised of Freunds Incomplete Adjuvant. The dose of peptide for the initial immunization is from about 1 to about 50,000 µg, generally 100-5,000 µg, for a 70 kg
20 patient. The initial administration of vaccine is followed by booster dosages at 4 weeks followed by evaluation of the magnitude of the immune response in the patient, by techniques that determine the presence of epitope-specific CTL populations in a PBMC sample. Additional booster doses are administered as required. The composition is found to be both safe and efficacious as a prophylaxis against HCV infection.

25 Alternatively, the polyepitopic peptide composition can be administered as a nucleic acid in accordance with methodologies known in the art and disclosed herein.

Example 14: Polyepitopic Vaccine Compositions Derived from Native HCV Sequences

A native HCV polyprotein sequence is screened, preferably using computer
30 algorithms defined for each class I and/or class II supermotif or motif, to identify “relatively short” regions of the polyprotein that comprise multiple epitopes and is preferably less in length than an entire native antigen. This relatively short sequence that contains multiple distinct, even overlapping, epitopes is selected and used to generate a minigene construct. The construct is engineered to express the peptide, which

corresponds to the native protein sequence. The "relatively short" peptide is generally less than 250 amino acids in length, often less than 100 amino acids in length, preferably less than 75 amino acids in length, and more preferably less than 50 amino acids in length. The protein sequence of the vaccine composition is selected because it has

5 maximal number of epitopes contained within the sequence, *i.e.*, it has a high concentration of epitopes. As noted herein, epitope motifs may be nested or overlapping (*i.e.*, frame shifted relative to one another). For example, with frame shifted overlapping epitopes, two 9-mer epitopes and one 10-mer epitope can be present in a 10 amino acid peptide. Such a vaccine composition is administered for therapeutic or prophylactic

10 purposes.

The vaccine composition will preferably include, for example, three CTL epitopes and at least one HTL epitope from an HCV antigen. This polyepitopic native sequence is administered either as a peptide or as a nucleic acid sequence which encodes the peptide. Alternatively, an analog can be made of this native sequence, whereby one or more of the

15 epitopes comprise substitutions that alter the cross-reactivity and/or binding affinity properties of the polyepitopic peptide.

The embodiment of this example provides for the possibility that an as yet undiscovered aspect of immune system processing will apply to the native nested sequence and thereby facilitate the production of therapeutic or prophylactic immune

20 response-inducing vaccine compositions. Additionally such an embodiment provides for the possibility of motif-bearing epitopes for an HLA makeup that is presently unknown. Furthermore, this embodiment (absent analogs) directs the immune response to multiple peptide sequences that are actually present in native HCV antigens thus avoiding the need to evaluate any junctional epitopes. Lastly, the embodiment provides an economy of

25 scale when producing nucleic acid vaccine compositions.

Related to this embodiment, computer programs can be derived in accordance with principles in the art, which identify in a target sequence, the greatest number of epitopes per sequence length.

30 Example 15. Polyepitopic Vaccine Compositions Directed To Multiple Diseases

The HCV peptide epitopes of the present invention are used in conjunction with peptide epitopes from target antigens related to one or more other diseases, to create a vaccine composition that is useful for the prevention or treatment of HCV as well as the

one or more other disease(s). Examples of the other diseases include, but are not limited to, HIV, and HBV.

For example, a polyepitopic peptide composition comprising multiple CTL and HTL epitopes that target greater than 98% of the population may be created for 5 administration to individuals at risk for both HCV and HIV infection. The composition can be provided as a single polypeptide that incorporates the multiple epitopes from the various disease-associated sources, or can be administered as a composition comprising one or more discrete epitopes.

10 **Example 16. Use of peptides to evaluate an immune response**

Peptides of the invention may be used to analyze an immune response for the presence of specific CTL or HTL populations directed to a prostate cancer-associated antigen. Such an analysis may be performed using multimeric complexes as described, e.g., by Ogg *et al.*, *Science* 279:2103-2106, 1998 and Greten *et al.*, *Proc. Natl. Acad. Sci. USA* 95:7568-7573, 1998. In the following example, peptides in accordance with the invention are used as a reagent for diagnostic or prognostic purposes, not as an immunogen.

In this example, highly sensitive human leukocyte antigen tetrameric complexes ("tetramers") are used for a cross-sectional analysis of, for example, HCV HLA-A*0201-specific CTL frequencies from HLA A*0201-positive individuals at different stages of disease or following immunization using an HCV peptide containing an A*0201 motif. 20 Tetrameric complexes are synthesized as described (Musey *et al.*, *N. Engl. J. Med.* 337:1267, 1997). Briefly, purified HLA heavy chain (A*0201 in this example) and β 2-microglobulin are synthesized by means of a prokaryotic expression system. The heavy 25 chain is modified by deletion of the transmembrane-cytosolic tail and COOH-terminal addition of a sequence containing a BirA enzymatic biotinylation site. The heavy chain, β 2-microglobulin, and peptide are refolded by dilution. The 45-kD refolded product is isolated by fast protein liquid chromatography and then biotinylated by BirA in the presence of biotin (Sigma, St. Louis, Missouri), adenosine 5' triphosphate and 30 magnesium. Streptavidin-phycoerythrin conjugate is added in a 1:4 molar ratio, and the tetrameric product is concentrated to 1 mg/ml. The resulting product is referred to as tetramer-phycoerythrin.

For the analysis of patient blood samples, approximately one million PBMCs are centrifuged at 300g for 5 minutes and resuspended in 50 µl of cold phosphate-buffered saline. Tri-color analysis is performed with the tetramer-phycoerythrin, along with anti-CD8-Tricolor, and anti-CD38. The PBMCs are incubated 5 with tetramer and antibodies on ice for 30 to 60 min and then washed twice before formaldehyde fixation. Gates are applied to contain >99.98% of control samples. Controls for the tetramers include both A*0201-negative individuals and A*0201-positive uninfected donors. The percentage of cells stained with the tetramer is then determined by flow cytometry. The results indicate the number of cells in the PBMC sample that 10 contain epitope-restricted CTLs, thereby readily indicating the extent of immune response to the HCV epitope, and thus the stage of HCV infection or exposure to a vaccine that elicits a protective or therapeutic response.

Example 17: Use of Peptide Epitopes to Evaluate Recall Responses

15 The peptide epitopes of the invention are used as reagents to evaluate T cell responses, such as acute or recall responses, in patients. Such an analysis may be performed on patients who have recovered from infection, who are chronically infected with HCV, or who have been vaccinated with an HCV vaccine.

20 For example, the class I restricted CTL response of persons who have been vaccinated may be analyzed. The vaccine may be any HCV vaccine. PBMC are collected from vaccinated individuals and HLA typed. Appropriate peptide epitopes of the invention that are preferably highly conserved and, optimally, bear supermotifs to provide cross-reactivity with multiple HLA supertype family members, are then used for analysis of samples derived from individuals who bear that HLA type.

25 PBMC from vaccinated individuals are separated on Ficoll-Histopaque density gradients (Sigma Chemical Co., St. Louis, MO), washed three times in HBSS (GIBCO Laboratories), resuspended in RPMI-1640 (GIBCO Laboratories) supplemented with L-glutamine (2mM), penicillin (50U/ml), streptomycin (50 µg/ml), and Hepes (10mM) containing 10% heat-inactivated human AB serum (complete RPMI) and plated using 30 microculture formats. A synthetic peptide comprising an epitope of the invention is added at 10 µg/ml to each well and HBV core 128-140 epitope is added at 1 µg/ml to each well as a source of T cell help during the first week of stimulation.

In the microculture format, 4 x 10⁵ PBMC are stimulated with peptide in 8 replicate cultures in 96-well round bottom plate in 100 µl/well of complete RPMI. On

days 3 and 10, 100 ml of complete RPMI and 20 U/ml final concentration of rIL-2 are added to each well. On day 7 the cultures are transferred into a 96-well flat-bottom plate and restimulated with peptide, rIL-2 and 10^5 irradiated (3,000 rad) autologous feeder cells. The cultures are tested for cytotoxic activity on day 14. A positive CTL response 5 requires two or more of the eight replicate cultures to display greater than 10% specific ^{51}Cr release, based on comparison with uninfected control subjects as previously described (Rehermann, *et al.*, *Nature Med.* 2:1104,1108, 1996; Rehermann *et al.*, *J. Clin. Invest.* 97:1655-1665, 1996; and Rehermann *et al.* *J. Clin. Invest.* 98:1432-1440, 1996).

Target cell lines are autologous and allogeneic EBV-transformed B-LCL that are 10 either purchased from the American Society for Histocompatibility and Immunogenetics (ASHI, Boston, MA) or established from the pool of patients as described (Guilhot, *et al.* *J. Virol.* 66:2670-2678, 1992).

Cytotoxicity assays are performed in the following manner. Target cells consist 15 of either allogeneic HLA-matched or autologous EBV-transformed B lymphoblastoid cell line that are incubated overnight with the synthetic peptide epitope of the invention at 10 μM , and labeled with 100 μCi of ^{51}Cr (Amersham Corp., Arlington Heights, IL) for 1 hour after which they are washed four times with HBSS.

Cytolytic activity is determined in a standard 4-h, split well ^{51}Cr release assay 20 using U-bottomed 96 well plates containing 3,000 targets/well. Stimulated PBMC are tested at effector/target (E/T) ratios of 20-50:1 on day 14. Percent cytotoxicity is determined from the formula: $100 \times [(\text{experimental release-spontaneous release})/\text{maximum release-spontaneous release}]]$. Maximum release is determined by lysis of targets by detergent (2% Triton X-100; Sigma Chemical Co., St. Louis, MO). Spontaneous release is <25% of maximum release for all experiments.

25 The results of such an analysis indicate the extent to which HLA-restricted CTL populations have been stimulated by previous exposure to HCV or an HCV vaccine.

The class II restricted HTL responses may also be analyzed. Purified PBMC are cultured in a 96-well flat bottom plate at a density of 1.5×10^5 cells/well and are stimulated with 10 $\mu\text{g}/\text{ml}$ synthetic peptide, whole antigen, or PHA. Cells are routinely plated in 30 replicates of 4-6 wells for each condition. After seven days of culture, the medium is removed and replaced with fresh medium containing 10U/ml IL-2. Two days later, 1 μCi ^3H -thymidine is added to each well and incubation is continued for an additional 18 hours. Cellular DNA is then harvested on glass fiber mats and analyzed for ^3H -thymidine

incorporation. Antigen-specific T cell proliferation is calculated as the ratio of ^3H -thymidine incorporation in the presence of antigen divided by the ^3H -thymidine incorporation in the absence of antigen.

5 Example 18: Induction Of Specific CTL Response In Humans

A human clinical trial for an immunogenic composition comprising CTL and HTL epitopes of the invention is set up as an IND Phase I, dose escalation study and carried out as a randomized, double-blind, placebo-controlled trial. Such a trial is designed, for example, as follows:

10 A total of about 27 subjects are enrolled and divided into 3 groups:

Group I: 3 subjects are injected with placebo and 6 subjects are injected with 5 μg of peptide composition;

Group II: 3 subjects are injected with placebo and 6 subjects are injected with 50 μg peptide composition;

15 Group III: 3 subjects are injected with placebo and 6 subjects are injected with 500 μg of peptide composition.

After 4 weeks following the first injection, all subjects receive a booster inoculation at the same dosage.

20 The endpoints measured in this study relate to the safety and tolerability of the peptide composition as well as its immunogenicity. Cellular immune responses to the peptide composition are an index of the intrinsic activity of this the peptide composition, and can therefore be viewed as a measure of biological efficacy. The following summarize the clinical and laboratory data that relate to safety and efficacy endpoints.

25 Safety: The incidence of adverse events is monitored in the placebo and drug treatment group and assessed in terms of degree and reversibility.

Evaluation of Vaccine Efficacy: For evaluation of vaccine efficacy, subjects are bled before and after injection. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

30 The vaccine is found to be both safe and efficacious.

Example 19: Phase II Trials In Patients Infected With HCV

Phase II trials are performed to study the effect of administering the CTL-HTL peptide compositions to patients having chronic HCV infection. The main objectives of

the trials are to determine an effective dose and regimen for inducing CTLs in chronically infected HCV patients, to establish the safety of inducing a CTL and HTL response in these patients, and to see to what extent activation of CTLs improves the clinical picture of chronically infected CTL patients, as manifested by a transient flare in alanine 5 aminotransferase (ALT), normalization of ALT, and reduction in HCV DNA. Such a study is designed, for example, as follows:

The studies are performed in multiple centers. The trial design is an open-label, uncontrolled, dose escalation protocol wherein the peptide composition is administered as a single dose followed six weeks later by a single booster shot of the same dose. The 10 dosages are 50, 500 and 5,000 micrograms per injection. Drug-associated adverse effects (severity and reversibility) are recorded.

There are three patient groupings. The first group is injected with 50 micrograms of the peptide composition and the second and third groups with 500 and 5,000 15 micrograms of peptide composition, respectively. The patients within each group range in age from 21-65, include both males and females, and represent diverse ethnic backgrounds. All of them are infected with HCV for over five years and are HIV, HBV and delta hepatitis virus (HDV) negative, but have positive levels of HCV antigen.

The magnitude and incidence of ALT flares and the levels of HCV DNA in the blood are monitored to assess the effects of administering the peptide compositions. The 20 levels of HCV DNA in the blood are an indirect indication of the progress of treatment. The vaccine composition is found to be both safe and efficacious in the treatment of chronic HCV infection.

Example 20. Induction of CTL Responses Using a Prime Boost Protocol

25 A prime boost protocol can also be used for the administration of the vaccine to humans. Such a vaccine regimen may include an initial administration of, for example, naked DNA followed by a boost using recombinant virus encoding the vaccine, or recombinant protein/polypeptide or a peptide mixture administered in an adjuvant.

For example, the initial immunization may be performed using an expression 30 vector, such as that constructed in Example 11, in the form of naked nucleic acid administered IM (or SC or ID) in the amounts of 0.5-5 mg at multiple sites. The nucleic acid (0.1 to 1000 μ g) can also be administered using a gene gun. Following an incubation period of 3-4 weeks, a booster dose is administered. The booster can, e.g., be recombinant fowlpox virus administered at a dose of $5\text{-}10^7$ to $5\text{x}10^9$ pfu. An alternative

recombinant virus, such as an MVA, canarypox, adenovirus, or adeno-associated virus, can also be used for the booster, or the polyepitopic protein or a mixture of the peptides can be administered. For evaluation of vaccine efficacy, patient blood samples will be obtained before immunization as well as at intervals following administration of the 5 initial vaccine and booster doses of the vaccine. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

Analysis of the results will indicate that a magnitude of response sufficient to 10 achieve protective immunity or to treat HCV infection is generated.

Example 21. Administration of Vaccine Compositions Using Dendritic Cells

Vaccines comprising peptide epitopes of the invention may be administered using dendritic cells. In this example, the peptide-pulsed dendritic cells can be administered to 15 a patient to stimulate a CTL response *in vivo*. In this method dendritic cells are isolated, expanded, and pulsed with a vaccine comprising peptide CTL and HTL epitopes of the invention. The dendritic cells are infused back into the patient to elicit CTL and HTL responses *in vivo*. The induced CTL and HTL then destroy (CTL) or facilitate destruction (HTL) of the specific target HCV-infected cells that bear the proteins from which the 20 epitopes in the vaccine are derived.

Alternatively, *Ex vivo* CTL or HTL responses to a particular tumor-associated antigen can be induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of antigen-presenting cells, such as dendritic cells, and the appropriate immunogenic peptides. After an 25 appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cells, *i.e.*, tumor cells.

30 Example 22: Alternative Method of Identifying Motif-Bearing Peptides

Another way of identifying motif-bearing peptides is to elute them from cells bearing defined MHC molecules. For example, EBV transformed B cell lines used for tissue typing, have been extensively characterized to determine which HLA molecules they express. In certain cases these cells express only a single type of HLA molecule.

These cells can then be infected with a pathogenic organism, *e.g.*, HCV, or transfected with nucleic acids that express the antigen of interest. Thereafter, peptides produced by endogenous antigen processing of peptides produced consequent to infection (or as a result of transfection) will bind be displayed on the cell surface. The peptides are then 5 eluted from the HLA molecules by exposure to mild acid conditions and their amino acid sequence determined, *e.g.*, by mass spectral analysis (*e.g.*, Kubo *et al.*, *J. Immunol.* 152:3913, 1994). Because, as disclosed herein, the majority of peptides that bind a particular HLA molecule are motif-bearing, this is an alternative modality for obtaining the motif-bearing peptides correlated with the particular HLA molecule expressed on the 10 cell.

Alternatively, cell lines that do not express any endogenous HLA molecules can be transfected with an expression construct encoding a single HLA allele. These cells may then be used as described, *i.e.*, they may be infected with a pathogenic organism or transfected with nucleic acid encoding an antigen of interest to isolate peptides 15 corresponding to the pathogen or antigen of interest that have been presented on the cell surface. Peptides obtained from such an analysis will bear motif(s) that correspond to binding to the single HLA allele that is expressed in the cell.

As appreciated by one in the art, one can perform a similar analysis on a cell bearing more than one HLA allele and subsequently determine peptides specific for each 20 HLA allele expressed. Moreover, one of skill would also recognize that means other than infection or transfection, such as loading with a protein antigen, can be used to provide a source of antigen to the cell.

The above examples are provided to illustrate the invention but not to limit its scope. For example, the human terminology for the Major Histocompatibility Complex, 25 namely HLA, is used throughout this document. It is to be appreciated that these principles can be extended to other species as well. Thus, other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, and patent application cited herein are hereby incorporated by reference for all purposes.

TABLE I

SUPERMOTIFS	POSITION	POSITION	POSITION
	2 (Primary Anchor)	3 (Primary Anchor)	C Terminus (Primary Anchor)
A1	T, I, L, V, M, S		F, W, Y
A2	L, I, V, M, A, T, Q		I, V, M, A, T, L
A3	V, S, M, A, T, L, I		R, K
A24	Y, F, W, I, V, L, M, T		F, I, Y, W, L, M
B7	P		V, I, L, F, M, W, Y, A
B27	R, H, K		F, Y, L, W, M, I, V, A
B44	E, D		F, W, L, I, M, V, A
B58	A, T, S		F, W, Y, L, I, V, M, A
B62	Q, L, I, V, M, P		F, W, Y, M, I, V, L, A
<hr/>			
MOTIFS			
A1	T, S, M		Y
A1		D, E, A, S	Y
A2.1	L, M, V, Q, I, A, T		V, L, I, M, A, T
A3	L, M, V, I, S, A, T, F, C, G, D		K, Y, R, H, F, A
A11	V, T, M, L, I, S, A, G, N, C, D, F		K, R, Y, H
A24	Y, F, W, M		F, L, I, W
A*3101	M, V, T, A, L, I, S		R, K
A*3301	M, V, A, L, F, I, S, T		R, K
A*6801	A, V, T, M, S, L, I		R, K
B*0702	P		L, M, F, W, Y, A, I, V
B*3501	P		L, M, F, W, Y, I, V, A
B51	P		L, I, V, F, W, Y, A, M
B*5301	P		I, M, F, W, Y, A, L, V
B*5401	P		A, T, I, V, L, M, F, W, Y

Bolded residues are preferred, italicized residues are less preferred: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

TABLE II

SUPERMOTIFS	POSITION					C-terminus
	1	2	3	4	5	
A1	1° Anchor T,I,L,V,M,S					1° Anchor F,W,Y
A2	1° Anchor L,I,V,M,A, T,Q					1° Anchor L,I,V,M,A,T
A3 preferred	1° Anchor V,S,M,A,T, L,I	Y,F,W (4/5)			Y,F,W (3/5)	1° Anchor R,K
deleterious	D,E (3/5); P (5/5)		D,E (4/5)			
A24	1° Anchor Y,F,W,I,V, L,M,T					1° Anchor F,I,Y,W,L,M
B7 preferred	F,W,Y (5/5) L,I,V,M (3/5)	1° Anchor P	F,W,Y (4/5)			
deleterious	D,E (3/5); P (5/5); G (4/5); A (3/5); Q,N (3/5)			D,E (3/5)	G (4/5)	1° Anchor V,I,L,F,M,W,Y,A
					Q,N (4/5)	D,E (4/5)
B27	1° Anchor R,H,K					1° Anchor F,Y,I,W,M,V,A
B44	1° Anchor E,D					1° Anchor F,W,Y,I,M,V,A
B58	1° Anchor A,T,S					1° Anchor F,W,Y,L,I,V,M,A
B62	1° Anchor Q,L,I,V,M, P					1° Anchor F,W,Y,M,I,V,L,A

MOTIFS		POSITION								
		1	2	3	4	5	6	7	8	C-terminus
A1 9-mer	preferred	G,F,Y,W	^{1°} Anchor S,T,M	D,E,A	Y,F,W		P	D,E,Q,N	Y,F,W	^{1°} Anchor Y
	deleterious	D,E		R,H,K,L,I,V M,P	A	G	A			
A1 9-mer	preferred	G,R,H,K	A,S,T,C,L,I V,M, ^{1°} Anchor D,E,A,S	G,S,T,C		A,S,T,C	L,I,V,M	D,E	^{1°} Anchor Y	
	deleterious	A	R,H,K,D,E, P,Y,F,W	D,E	P,Q,N	R,H,K	P,G	G,P		

		POSITION									
		1	2	3	4	5	6	7	8	9	
A.1 10-mer		Y,F,W	Y,F,W	1°Anchor S,T,M	D,E,A,Q,N	A	Y,F,W,Q,N	P,A,S,T,C	G,D,E	P	1°Anchor Y
deleterious		G,P		R,H,K,G,L,I V,M	D,E	R,H,K	Q,N,A W	R,H,K,Y,F, W	R,H,K	A	C-terminus
A.1 10-mer		Y,F,W	Y,F,W	S,T,C,L,I,V M	1°Anchor D,E,A,S	A	Y,F,W	P,G	G	Y,F,W	1°Anchor Y
deleterious		R,H,K		R,H,K,D,E, P,Y,F,W		P	G	P,R,H,K	Q,N		
A.2.1 9-mer		Y,F,W	Y,F,W	1°Anchor L,M,I,V,Q, A,T	Y,F,W	S,T,C	Y,F,W	A	P	1°Anchor V,L,I,M,A,T	
deleterious		D,E,P		D,E,R,K,H		R,K,H	D,E,R,K,H				
A.2.1 10-mer		A,Y,F,W	A,Y,F,W	1°Anchor L,M,I,V,Q, A,T	L,V,I,M	G	G	F,Y,W, L,V,I,M		1°Anchor V,L,I,M,A,T	
deleterious		D,E,P		D,E	R,K,H,A	P	R,K,H	D,E,R, K,H	R,K,H		

		POSITION									
		1	2	3	4	5	6	7	8	9	C-terminus
A3	preferred	R,H,K	^{1°} Anchor L,M,V,I,S, A,T,F,C,G D	Y,F,W	P,R,H,K,Y, F,W	A	Y,F,W		P		C-terminus ^{1°} Anchor K,Y,R,H,F,A
	deleterious	D,E,P		D,E							
A11	preferred	A	^{1°} Anchor V,T,L,M,I, S,A,G,N,C, D,F	Y,F,W	Y,F,W	A	Y,F,W	Y,FW	P		^{1°} Anchor K,R,Y,H
	deleterious	D,E,P						A	G		
A24 9-mer	preferred	Y,F,W,R,H,K	^{1°} Anchor Y,F,W,M		S,T,C			Y,F,W	Y,F,W		^{1°} Anchor F,I,I,W
	deleterious	D,E,G	D,E	G	Q,N,P	D,E,R,H,K	G				
A24 10-mer	preferred		^{1°} Anchor Y,F,W,M	P	Y,F,W,P			P			^{1°} Anchor F,L,I,W
	deleterious		G,D,E	Q,N	R,H,K	D,E	A	Q,N	D,E,A		
A310i	preferred	R,H,K	^{1°} Anchor M,V,T,A,L, I,S	Y,F,W	P		Y,F,W	Y,F,W	A,P		^{1°} Anchor R,K
	deleterious	D,E,P	D,E	A,D,E	D,E	D,E	D,E	D,E			

		POSITION								C-terminus	
		1	2	3	4	5	6	7	8	or	
A3301	preferred									A, Y, F, W	C-terminus
		1°Anchor M,V,A,L,F, I,S,T		Y,F,W						1°Anchor R,K	1°Anchor R,K
deleterious	G,P										
A6801	preferred	Y,F,W,S,T,C		1°Anchor A,V,T,M,S, L,I				Y,F,W,L,I, V,M		P	1°Anchor R,K
deleterious	G,P			D,E,G				R,H,K		A	
B0702	preferred	R,H,K,F,W,Y		1°Anchor P	R,H,K			R,H,K		P,A	1°Anchor L,M,F,W,Y,A, I,V
deleterious	D,E,Q,N,P			D,E,P	D,E			D,E	G,D,E	Q,N	D,E
B3501	preferred	F,W,Y,L,I,V,M		1°Anchor P	F,W,Y			F,W,Y		F,W,Y	1°Anchor L,M,F,W,Y,I, V,A
deleterious	A,G,P							G		G	

		POSITION									
		1	2	3	4	5	6	7	8	9	C-terminus
B51	preferred	L,I,V,M,F,W,Y	^{1°} Anchor P	F,W,Y	S,T,C	F,W,Y	G		F,W,Y	F,W,Y	C-terminus 1°Anchor L,I,V,F,W, Y,A,M
	deleterious	A,G,P,D,E,R,H,K, S,T,C			D,E	G		D,E,Q,N	G,D,E		
B5301	preferred	L,I,V,M,F,W,Y	^{1°} Anchor P	F,W,Y	S,T,C	F,W,Y		L,I,V,M,F, W,Y	F,W,Y	F,W,Y	1°Anchor I,M,F,W,Y, A,L,V
	deleterious	A,G,P,Q,N				G		R,H,K,Q,N	D,E		
B5401	preferred	F,W,Y	^{1°} Anchor P	F,W,Y,L,I,V M		L,I,V,M	A,L,I,V,M	F,W,Y,A,P		1°Anchor A,T,I,V,L, M,F,W,Y	
	deleterious	G,P,Q,N,D,E		G,D,E,S,T,C		R,H,K,D,E	D,E	Q,N,D,G,E	D,E		

Italicized residues indicate less preferred or "tolerated" residues.
The information in Table II is specific for 9-mers unless otherwise specified.

Table III

MOTIFS		POSITION							
	1 ^o anchor 1	2	3	4	5	6 ^o anchor 6	7	8	9
DR4 preferred	F, M, Y, L, I, <i>V, W</i>	M	T	I	V, S, T, C, P, A, <i>L, I, M</i>	M, H,			M, H
deleterious				W,		R,			W, D, E
DR1 preferred	M, F, L, I, V, <i>W, Y</i>	C	C, H	F, D	P, A, M, Q	V, M, A, T, S, P, <i>L, I, C</i>	M,		A, V, M
deleterious									
DR7 preferred	M, F, L, I, V, <i>W, Y</i>	M	W	A		I, V, M, S, A, C, <i>T, P, L</i>	M		I, V
deleterious									
DR Supermotif	M, F, L, I, V, <i>W, Y</i>	C,	C,			G, R, D	N		G
DR3 MOTIFS		1 ^o anchor 1	2	3	4 ^o anchor 4	5	6 ^o anchor 6		
motif a preferred	L, I, V, M, F, <i>Y</i>					D			
motif b preferred	L, I, V, M, F, <i>A, Y</i>				D, N, Q, E, <i>S, T</i>	K, R, H			

Italicized residues indicate less preferred or "tolerated" residues.

Table IV: HLA Class I Standard Peptide Binding Affinity.

ALLEL	STANDARD PEPTIDE	SEQUENCE (SEQ ID NO:)	STANDARD BINDING AFFINITY (nM)
A*0101	944.02	YLEPAIAKY	25
A*0201	941.01	FLPSDYFPSV	5.0
A*0202	941.01	FLPSDYFPSV	4.3
A*0203	941.01	FLPSDYFPSV	10
A*0205	941.01	FLPSDYFPSV	4.3
A*0206	941.01	FLPSDYFPSV	3.7
A*0207	941.01	FLPSDYFPSV	23
A*6802	1072.34	YVIKVSARV	8.0
A*0301	941.12	KVFPYALINK	11
A*1101	940.06	AVDLYHFLK	6.0
A*3101	941.12	KVFPYALINK	18
A*3301	1083.02	STLPETYVVRR	29
A*6801	941.12	KVFPYALINK	8.0
A*2402	979.02	AYIDNYNKF	12
B*0702	1075.23	APRTLVYLL	5.5
B*3501	1021.05	FPFKYAAAF	7.2
B51	1021.05	FPFKYAAAF	5.5
B*5301	1021.05	FPFKYAAAF	9.3
B*5401	1021.05	FPFKYAAAF	10

Table V. HLA Class II Standard Peptide Binding Affinity.

Allele	Nomenclature	Standard Peptide	Sequence (SEQ ID NO:)	Binding Affinity (nM)
DRB1*0101	DR1	515.01	PKYVKQNTLKLAT	5.0
DRB1*0301	DR3	829.02	YKTIAFDEEARR	300
DRB1*0401	DR4w4	515.01	PKYVKQNTLKLAT	45
DRB1*0404	DR4w14	717.01	YARFQSQTTLKQKT	50
DRB1*0405	DR4w15	717.01	YARFQSQTTLKQKT	38
DRB1*0701	DR7	553.01	QYIKANSKFIGITE	25
DRB1*0802	DR8w2	553.01	QYIKANSKFIGITE	49
DRB1*0803	DR8w3	553.01	QYIKANSKFIGITE	1600
DRB1*0901	DR9	553.01	QYIKANSKFIGITE	75
DRB1*1101	DR5w11	553.01	QYIKANSKFIGITE	20
DRB1*1201	DR5w12	1200.05	EALIHQLKINPYVLS	298
DRB1*1302	DR6w19	650.22	QYIKANAKFIGITE	3.5
DRB1*1501	DR2w2 β 1	507.02	GRTQDENPVVHFFKNIV TPRTPPP	9.1
DRB3*0101	DR52a	511	NGQIGNDPNRDIL	470
DRB4*0101	DRw53	717.01	YARFQSQTTLKQKT	58
DRB5*0101	DR2w2 β 2	553.01	QYIKANSKFIGITE	20

Table VI

HLA-supertype	Allele-specific HLA-supertype members		Predicted ^b
	Verified ^a		
A1	A*0101, A*2501, A*2601, A*2602, A*3201		A*0102, A*2604, A*3601, A*4301, A*8001
A2	A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, A*6901		A*0208, A*0210, A*0211, A*0212, A*0213
A3	A*0301, A*1101, A*3101, A*3301, A*6801		A*0302, A*1102, A*2603, A*3302, A*3303, A*3401, A*3402, A*6601, A*6602, A*7401
A24	A*2301, A*2402, A*3001		A*2403, A*2404, A*3002, A*3003
B7	B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3503, B*3504, B*3505, B*3505, B*3506, B*3507, B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, B*7801		B*1511, B*4201, B*3901
B27	B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*2701, B*2707, B*2708, B*3802, B*3903, B*3904, B*3801, B*3901, B*3902, B*7301		B*3905, B*4801, B*4802, B*1510, B*1518, B*1503
B44	B*1801, B*1802, B*3701, B*4402, B*4403, B*4404, B*4001, B*4002, B*4006		B*4101, B*4501, B*4701, B*4901, B*5001
B58	B*5701, B*5702, B*5801, B*5802, B*1516, B*1517		
B62	B*1501, B*1502, B*1513, B*5201	B*1301, B*1302, B*1504, B*1505, B*1506, B*1507, B*1515, B*1520, B*1521, B*1512, B*1514, B*1510	

a. Verified alleles include alleles whose specificity has been determined by pool sequencing analysis, peptide binding assays, or by analysis of the sequences of CTL epitopes.

b. Predicted alleles are alleles whose specificity is predicted on the basis of B and F pocket structure to overlap with the supertype specificity.

Table VII HCV_A01_Super_Motif_With_Binding_Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'0101
ATGNLPGCSF	165	10	1.3	93	
ATLGFCAF	1285	8	1.4	100	
AVGWNHHLAF	1917	11	1.4	100	
CTCGSSLY	1128	9	1.1	78	0.3700
CTRGAVAKAVDF	1190	11	1.1	79	
CTWANSTGFF	555	9	1.1	79	
CVTOTVDF	1462	8	1.2	88	
CVLAVTSRV	1657	9	1.2	88	
ETTMPSVPVF	1207	9	1.2	88	
FSDTTRCF	2870	0	1.1	79	
FTEAMTRY	2792	8	1.4	100	
FTGLTHIDAF	1567	11	1.3	93	
GIPWQDHLF	1552	11	1.2	88	
GLSFSLHSY	2921	10	1.1	78	0.0028
GLTHIDAF	1569	9	1.3	93	
GSGYGFY	2641	8	1.1	79	
GTFINAY	2063	8	1.1	79	
GVAGALVAF	1653	9	1.2	88	
GVAYDF	1183	8	1.1	79	
GVLAALAY	1670	9	1.2	88	
GVRCERWALY	2619	11	1.4	100	
GVRLLEDGARY	154	11	1.2	88	
HUHQINVOY	696	11	1.1	79	
HMNPNSGIOY	1769	11	1.3	93	
HNGPREGAVOW	1910	11	1.1	79	
IMAKNEVF	2581	8	1.2	88	
ITYSTYCKF	1206	8	1.2	88	
IVDQVLY	701	8	1.2	88	
KSTKVPAAY	1241	9	1.2	88	
KVIDLTCGF	121	10	1.2	88	
LIEANLW	2235	8	1.2	88	
LINTNGSV	414	8	1.1	79	
LLAPITAY	1030	8	1.4	100	
LLPMLGGW	1812	9	1.2	88	
LLSPRGSRPSW	97	11	1.1	79	0.8100
LSASLHSY	2922	9	1.1	79	
LSFRGSRPSW	88	10	1.1	79	
LTCGFDALMAY	126	11	1.2	88	
LTHIDAF	1570	8	1.3	93	
LYDILQY	1853	8	1.1	79	
MILMTHFF	2878	8	1.2	88	
NIVDQVLY	700	9	1.2	88	0.0980
NLPGCSFSIF	168	10	1.3	93	
NTCYTOTVDF	1400	10	1.2	88	
NTNRPPODWF	14	11	1.1	78	

HCV_A01_Super_Motif_With_Blinding_Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	λ 0101
INDQDVGW	1108	0	11	79	
PITYSTYGF	1295	10	11	79	
PMGFSYDTRCF	2687	11	11	79	
PSVAAATLGF	1281	9	14	100	
PTLHGPTFLY	1621	11	11	79	
PVCOCHLFLF	1554	9	12	86	
PVCOQHLEFW	1554	10	12	86	
QTVQFSQDQPTF	1485	11	12	86	
RFLGSLAF	2910	6	12	86	
FLLAPITAY	1029	9	12	86	
RMAWDMMMANW	317	10	12	86	
RMILMTHF	2875	0	12	86	
RMLMTRIFF	2875	9	12	86	
RVCEKMAVY	2621	9	14	100	
RVLEDGVNV	1556	9	12	86	
STKVPAAY	1242	8	12	86	
SVAAATLGF	1262	8	14	100	
SVAATLGFQAY	1262	11	14	100	
TTIMAKNEVF	2680	9	11	79	
TLHGPTFLY	1622	10	11	79	
TLFNLQGGW	1811	10	12	86	
TTIMAKNEVF	2509	10	11	79	
TTMRSIPVF	1208	8	12	86	
TVDFSLQDPTF	1406	10	12	86	
VIDLTLCGF	122	9	12	86	
VLAALAY	1671	8	12	86	
VLEDGVNV	167	8	12	86	
VLDVILGY	1052	9	11	79	
VMGSSYGF	2639	6	11	79	
VMGSSYGFQY	2639	10	11	79	
VMNRLIAF	1920	6	14	100	
YSPGQRFV	2648	9	11	79	
YTNVQDQDVGW	1106	11	11	79	
YVGDLGGSVF	276	10	12	86	

2

79

Table VIII
HCV A02 Super Matrix with Binding Information

Conservancy	Freq.	Position	Sequence	A'0201	A'0202	A'0203	A'0206	A'6002
93	13	1904	AAILRNHV					
86	12	1673	AAIAAYCL					
79	11	1250	AAQGYKVL					
79	11	1250	AAQGYKVLV					
79	11	1250	AAQGYKVLV					
79	11	147	AAIAIAHGV					
79	11	147	AAIAIAHGVV					
100	14	1264	AATLGFGA					
93	13	1264	AATLGFGAYM					
86	12	1187	AVCTRGV					
79	11	1187	AVCTRGV					
79	11	1187	AVCTRGVAKA					
93	13	1880	AILSPGAL					
86	12	1880	AILSPGALV					
86	12	1880	AILSPGALV					
100	14	150	ALAHGVHV					
100	14	150	ALAHGVRV					
66	12	1737	ALGLDQTA					
86	12	688	ALSTGQHL					
79	11	1896	ALVGIVCA					
79	11	1896	ALVGIVVCAA					
79	11	1896	ALVGIVVCAA					
66	12	1602	ADAPPSSWQDM					
79	11	1251	AGQGYKVL					
79	11	1251	AGQGYKVL					
86	12	77	AGPGTPWNL					
93	13	1285	ATLGFGAYM					
79	11	1354	ATPGQSVT					
79	11	1596	ATVCARAQA					
100	14	1419	AVAYYRGL					
100	14	1419	AVAYYRGLDV					
79	11	1188	AVCTRGV					
79	11	1188	AVCTRGVAKA					
79	11	1188	AVCTRGVAKV					
100	14	1917	AVGWMNRL					
100	14	1917	AVGWMNRL					
93	13	1917	CAILRRIHV					
79	11	1530	CAYEELTPA					
86	12	2841	CLRLGVVPL					
86	12	738	CLVMMILL					
79	11	1653	CMASDLEV					

IICV A01 Super Moll with Binding Information

Conservancy	Freq.	Position	Sequence	A'0201	A'0202	A'0203	A'0206	A'6802
				0.0067				
79	11	1653	CMSADLEWV					
79	11	1653	CMSADLEWVT					
79	11	1128	CTCGSSSDL					
79	11	1128	CTCGSSDYL					
78	11	1128	CTCGSSDYLV					
78	11	1128	CTRGVAKA					
79	11	1190	CTRGVAKAV					
79	11	1190	CTWMNSTGFT					
79	11	555	CTYTYDFSL					
86	12	1462	DAGGAWYEL					
79	11	1527	DAHFLSOT					
100	14	1574	DILAGYGA					
86	12	1655	DILAGYGAGV					
79	11	1655	DILAGYGAGVA					
79	11	1655	DLGGSVFL					
86	12	279	DLGGSVFLV					
79	11	279	DLEVNTST					
86	12	1657	DLEVNTSTWV					
86	12	1657	DLEVNTSTWVL					
86	12	1657	-DLGVRVGEKM-					
93	13	2617	DLGVRVCEKMA					
79	11	132	DLMGYPL					
79	11	132	DLMGYPLV					
79	11	132	DLSDGSWST					
79	11	2412	DLSDGSWSTV					
79	11	2412	DVNLLPA					
79	11	1883	DVNLLPAI					
79	11	1883	DVNLLPAIL					
79	11	2772	DWVICESA					
86	12	1134	DLYLYTRHA					
86	12	1134	DLYLYTRHADV					
86	12	321	DMMMNWSPT					
86	12	994	DOAETAGA					
86	12	994	DOAETAGAII					
88	12	1339	DOAETAGARL					
86	12	1339	DTAACGDI					
86	12	124	DTACGDI					
86	12	124	DTTCGFADL					
86	12	124	DTTCGFADLM					
93	13	2673	DTTCFADST					

HCY AND Super Motif with Binding Information

Conservancy	Fseq.	Position	Sequence	A'0201	A'0202	A'0203	A'0206	A'0207
93	13	2673	DTTCFDFSTV					
93	13	2673	DTTCFDFSTV					
86	12	21	DYKPGEGCQI	0.0001				
86	12	21	DYKPGEGCQI	0.0001				
79	11	750	DVKPGGGQIV					
			EALENLV					
100	14	2794	EAMTRYSA					
86	12	2237	EANLWROEM					
93	13	1377	EIPFGKIA	0.0001				
93	13	1377	EIPFGKIA	0.0001				
100	14	2614	ELTSCSSANV					
78	11	666	ELSPILLST					
79	11	666	ELSPILLSTT	0.00003				
66	12	2245	EMGGNITRV					
86	12	1731	EDFKQKAL					
86	12	1731	EQFKQKALGL					
66	12	1731	EQFKQKALGL					
66	12	1342	ETAGARLV					
86	12	1342	ETAGARLV					
86	12	1342	ETAGARLV					
86	12	1342	ETAGARLVLA					
66	12	1207	ETMRSPV					
66	12	1207	ETMRSPV					
86	12	1659	EVVTSWV	0.0001				
86	12	1659	EVVTSWV	0.0001				
86	12	1659	EVVTSWVLY	0.0004				
93	13	130	FADLMGYI					
79	11	130	FADLMGYIPLV					
79	11	130	FADLMGYIPLV					
100	14	1927	FASRGNIV					
86	12	1927	FASRGNIVSP					
100	14	1773	FISGIOYL	0.1000				
100	14	1773	FISGIOYL	0.1000				
100	14	1773	FISGIOYLAGL					
79	11	1304	FLAGGCGGGAA					
86	12	177	FLLALCCG	0.0048				
86	12	177	FLLALCCG	0.0048				
93	13	728	FLLADARV					
86	12	1228	FQVNLHHA					
86	12	1228	FQVNLHAPT					
78	11	2646	FQYSPGQRV					
100	14	2782	FTEAMTRYSA					
93	13	1587	FTGLTHIDA					

HCV A'02 Super-Motif with Binding Information

Conservancy	Freq.	Pasilian	Sequence	A'0201	A'0202	A'0203	A'0206	A'6802
93	13	512	FTPSPVVV					
93	13	512	FTPSPVVGT					
93	13	512	FTSPVVGTT					
79	11	684	FTLPALEST					
79	11	684	FTLPALETL					
79	11	146	GARALAHGV					
86	12	992	GADTAACGDI					
86	12	992	GADTAACGDI					
86	12	1861	GAGVAGAL					
86	12	1861	GAGVAGALV					
86	12	1861	GAGVAGALVA					
88	12	350	GATWGVLA					
79	11	1895	GALWGV					
79	11	1895	GALWGVVCA					
79	11	1895	GALWGVVCAA					
86	12	1345	GARLWVLA					
79	11	1345	GARLWVLAT					
79	11	1345	GARLWVLATAT					
100	14	1916	GAVQWMMRL					
100	14	1916	GAVQWMMRLU					
100	14	1916	GAVQWMMRLIA					
100	14	1333	GIGTYLDDA					
100	14	1333	GIGTVDOAET					
100	14	1776	GICYLAGL					
100	14	1776	GIQYLAGLST					
100	14	1776	GIVDVSPI					
79	11	1426	GLPQCDL					
83	13	1552	GLPQCDL					
79	11	968	GLPDLAYA					
79	11	968	GLRDLAVAV					
100	14	1782	GLSLPGNPA					
79	11	1782	GLSLPGNPAI					
93	13	1589	GLTIIIDAHFL					
93	13	26	GCTGCCYCL					
93	13	26	GONGGMYL					
79	11	2063	GTFFPINAYT					
79	11	2063	GTFFPINAYTT					
100	14	1335	GTVLDQAET					
100	14	1335	GTVLDQAETA					
86	12	1863	GYAGALVA					
		1081	GYCWTWYHGA					
		78						

HCV AND Super Motif Binding Information

Conservancy	Freq.	Position	Sequence	A'0201	A'0202	A'0203	A'0208	A'6802
8.8	12	1870	GVLAAALAA					
8.6	12	1670	GVLAAAYCL					
7.9	11	161	GVNATGNL					0.0001
8.6	12	45	GVRAATKT					
10.0	14	2618	GVRYCEKM					
10.0	14	2619	GVRYCEKMA					
10.0	14	2619	GVRYCEKMA	0.0002				
8.3	13	154	GVRYEDGY	0.0001				
7.9	11	1900	GVYCAAIL					
10.0	14	1234	HAPTSGKST					
10.0	14	1572	HDAHFLSQT					
8.6	12	696	HUIONNOV					
7.9	11	1719	HUYIEQGM					
9.3	13	1769	HWMNFISGI					
7.9	11	698	HNONNOVYL					
7.9	11	222	HPPGCYPCV					
8.6	12	2855	HTPVNSWL					
8.6	12	2855	HTPVNSWLGN					
7.9	11	1910	HYPGEEGA					
7.9	11	1910	HYSPGEAV					
8.6	12	1933	HSPTHYY					
10.0	14	1925	IAFASRGNHV					
7.9	11	1856	ILAGYGAGV					
7.9	11	1856	ILAGYGAGVA					
8.6	12	1816	ILGGWVAAQL					
8.6	12	1816	ILGGWVAAQLA					
8.6	12	1331	ILGGTIVL					
8.6	12	1331	ILGGTIVLDOA					
9.3	13	1891	ILSPGALVY					
9.3	13	1891	ILSPGALVY					
9.3	13	1891	ILSPGALVGY					
7.9	11	2591	IMARNEVFCV					
10.0	14	1777	IQYLAGLST					
10.0	14	1777	IQYLAGLSTL					
8.6	12	2250	ITVSESENKV					
8.6	12	2250	ITVSESENKVW					
10.0	14	2816	ITSCSSNNV					
10.0	14	2816	ITSCSSNNVSV					
10.0	14	2816	ITSCSSNNVSA					
8.6	12	909	ITWGADTA					
8.6	12	909	ITWGADTA					

HCV_A02_Super_Motif_With_Bindline_Information

Conservancy	Freq.	Position	Sequence	A'0201	A'0202	A'0203	A'0206	A'6802.
79	11	1296	ITYSTYGGKEL					
79	11	1296	ITYSTYGGKFLA					
79	11	2813	NFPDLGVV	0.0016				
79	11	2813	IVFPDLGVV					
93	13	30	IVGGYLL					
86	12	1738	KALGILLOT					
86	12	1738	KALGILLOTA					
86	12	2625	KMAYDV					
86	12	2625	KMAYDVL					
86	12	1734	KOKALGILL					
86	12	1734	KOKALGILLOT					
86	12	1734	KOKALGILLOTA					
86	12	121	KVIDLTCGFA	0.0048				
100	14	1255	KVLYVNPSCV					
100	14	1255	KVLYVNPSCVA					
100	14	1255	KVLYVNPSCVAA					
79	11	1244	KVPVAYAA	0.0011				
86	12	1872	LAALAYCL					
79	11	1305	LADEGCCSGA					
86	12	1729	LAEDFKQKA					
86	12	1729	LAEDFKQKAL					
79	11	1657	LAGYGAJV					
79	11	1857	LAGYGAJVVA					
79	11	1857	LAGYGAJVAGA					
100	14	151	LAHGIVRL					
86	12	179	LALLSCLT					
79	11	972	LAWAVEPV					
100	14	1924	LIAFASRGNHV	0.0004				
100	14	2815	LITSCSSNV					
100	14	2815	LITSCSSNVV					
79	11	2812	LNFPDLGVV	0.0002				
79	11	2812	LALLSCL					
86	12	178	LLALLSCLT					
86	12	178	LLFLLLADA	0.0230	0.0150	0.0220	0.0011	0.0130
100	14	726	LLFLLLADAV					
93	13	726	LLFLLLADAVV	1.2000	0.0380	3.1000	0.1800	1.2000
86	12	1812	LLFLNLGGWV					
86	12	1812	LLFLNLGGWVA					
93	13	728	LLLADAV	0.0061				
93	13	1887	LLPAILSPGA					
93	13	1887	LLPAILSPGAL					
93	13	38	LLPRRGSPFL	0.0025				
93	13	36	LLPRRGSPFLV					

HCV A02 Super Motif with Blueline Information

Conservancy	Freq.	Position	Sequence	A'0201	A'0202	A'0203	A'0208	A'0802
86	12	2240	LLWADENGGN					
93	13	1629	LLYRLGAV					
79	11	133	LMGIPFLV					
79	11	133	LMGYPLVGA					
86	12	2761	LOCCTMLV					
86	12	126	LTCGFADL					
86	12	126	LTCGFADLM					
100	14	2180	LTDPHIT					
100	14	2180	LTDPHITA					
86	12	1052	LTGRDKNOV					
93	13	1570	LTMIONHFL					
93	13	2176	LTSMLTOPSHI					
79	11	2738	LTTSCGNT					
79	11	2738	LTTSCGNTL					
79	11	2738	LTTSCGNTLT					
86	12	1591	LVAYQATV					
86	12	1591	LVAYQATYCA					
79	11	1863	LVDIAGIGA					
86	12	1867	LVGGVIAA					
86	12	1867	LVGGVLAAL					
86	12	1867	LVGGVLAALA					
86	12	1867	LVGGVLAALAA					
100	14	1257	LVNLPSSVA					
100	14	1257	LVNLPSSVAAT					
100	14	1257	LVNLPSSVAATL					
100	14	1257	LVNLPNAI					
79	11	1684	LVNLLPAIL					
79	11	1684	LVTRHADV					
86	12	1137	LVTRHADVI					
79	11	1137	LVTRHADVIPV					
79	11	1137	LVVGWCA					
79	11	1697	LVVGWCAA					
79	11	1697	LVVGWVCAA					
79	11	1697	LVVGWVCAAII					
79	11	2773	LVVICESA					
66	12	1348	LVVLATAT					
86	12	2582	MAKNEVFCV					
100	14	2179	MLTOPSHI					
100	14	2179	MLTDPHIT					
100	14	2179	MLTDPHSITA					
93	13	322	MMMMNWSP					

HCV AND Human Monocyte Binding Information

Conservancy	Freq.	Position	Sequence	A' 0201	A' 0202	A' 0203	A' 0206	A' 6802
9.3	13	1416	NAVAYYAGL					
9.3	13	1418	NAVAYYNGLOV					
8.8	12	2068	NAYTTGPCT					
8.6	12	1815	NILGGWVA					
8.6	12	1815	NILGGWVAAL					
8.6	13	1282	NIRTGVRT					
9.3	11	1282	NIRTGVATI					
7.9	11	1282	NIRTGVNTT					
7.9	11	1282	NIRTGVTTT					
6.6	12	2249	NITRVESENRY					
8.8	12	700	NIVDOYVL					
8.8	12	116	NLGKVIOT					
8.8	12	116	NLGKVIDTL					
8.6	12	116	NLGKVIDLT					
9.3	13	1888	NLLPAISPGA					
8.8	12	2239	NLWRAQEM					
9.3	13	168	NLPGCSFSI					
9.3	13	168	NLPGCSFSI					
8.6	12	1480	NTCVTQTV					
9.3	13	416	NTCSWHL					
8.6	12	14	NTNFRPODV					
9.3	13	1889	PAILSP'GA					
9.3	13	1889	PAILSP'GAL					
8.6	12	1889	PAILSP'GALV					
8.8	12	1889	PAILSP'GALV					
8.6	12	888	PALSTGLI					
8.6	12	668	PALSTGUHL					
7.9	11	2609	PARIUVFFL					
7.9	11	2088	PINAYTTGPCT					
7.9	11	1285	PITYSTYKGFL					
9.3	13	2403	PLEGEPEPQL					
7.9	11	143	PLGGAARA					
7.9	11	143	PLGGAARAL					
7.9	11	143	PLGGAAAL					
9.3	13	1628	PLYYRGLA					
9.3	13	1628	PLYYRGLAV					
7.9	11	2667	PNGFSDT					
7.9	11	2807	POPEYDEL					
7.9	11	2807	POPEYOLEU					
7.9	11	2807	POPEYOLEUT					
9.3	13	7	PORKTRKT					

HCV A02 Super Motif with Biotin Information

Conservancy	Freq.	Position	Sequence	A'0201	A'0202	A'0203	A'0206	A'6802
86	12	109	PTDPARIANL					
79	11	1473	PTFETTT					
79	11	1473	PTFTETTT					
100	14	1236	PTGSGKST					
93	13	1236	PTGSGKSTKV					
86	12	1936	PTHVVPESDA					
86	12	1936	PTIIMYFESDAA					
79	11	1621	PTLHQPTPL					
79	11	1621	PTLHQPTPLL					
79	11	2070	PTLWAHMI					
79	11	2870	PTLWAHMIIL					
79	11	2870	PTLWAHMIILM					
79	11	2870	PTLWAHMIILMT					
100	14	1626	PTPLLYAL					
93	13	1826	PTPLLYALGA					
93	13	1826	PTPLLYRLGAV					
100	14	2857	PVNSVGNI					
100	14	2857	PVNSVGNII					
86	12	2857	PVNSWLGNIIM					
79	11	2318	PWVHGCPPL					
93	13	508	PYCGFTPSPV					
93	13	508	PYCGFTPSVW					
86	12	1340	QAETAGAFL					
86	12	1340	QAETAGAHLV					
86	12	1340	DAETAGAHLW					
80	12	1603	QAPPSSWQDM					
93	13	1595	QATVCAIA					
79	11	1595	QATVCAIAQ					
83	13	29	QVGGVYL					
83	13	29	QVGGVYL					
86	12	336	QLRPOA					
86	12	2184	QLPCEPFDV					
79	11	2210	QLSASFSLKA					
79	11	2210	QLSASFSLKAT					
86	12	1465	QTVGFSCLFT					
86	12	1229	QVAKLHAPT					
86	12	1186	RAAVCTRGV					
79	11	1186	RAAVCTRGVA					
100	14	149	RALAHGVAV					
100	14	149	RALAHGVAVL					
86	12	2733	RASGVLT					
		43	RLGVATRKT					
		11						
		79						

ICV A01 Similar Motif with Binding Information

Conservancy	Freq.	Position	Sequence	A'0201	A'0202	A'0203	A'0206	A'6802
79	11	2916	FLHGSASFSL	0.0280	0.0055	0.0180	0.0002	0.0032
79	11	2611	RLVFFDL	0.0890	0.0110	1.0000	0.0100	0.0050
79	11	2611	RLVFPDYGIV					
79	11	1618	PLKPTLHGP					
86	12	1029	RLLAPITA					
86	12	1347	RLVVLATA					
86	12	1347	RLVVLATAT					
100	14	619	RLWHPCT					
86	12	3117	RMAMWDMMM					
93	13	635	AMVYGGAEHNL					
85	12	2243	RQENGGNI					
88	12	2243	RQEMGGNIT					
86	12	2243	RGEMGNITRY					
79	11	1284	RTGVRITIT					
79	11	1284	RTGVRITTT					
100	14	2621	RVCEKMAVL					
86	12	2621	RVCEKMAVLVD					
86	12	2252	RVESENKV					
86	12	2252	RVESENKVV					
79	11	2100	RVGDRIV					
86	12	156	RVLEDGVNYA					
86	12	156	RVLEDGVNYAT					
86	12	2833	RVYLTDQPT					
79	11	1855	SADLEVIT					
79	11	1655	SADLEVITST					
79	11	2212	SAPSLKAT					
79	11	2212	SAPSLKATCT					
93	13	2207	SASQSLASPSL					
100	14	175	SIFLLALL					
88	12	175	SIFLLALLSCL					
100	14	1470	SLOPFTI					
86	12	1470	SLOPFTIET					
79	11	1470	SLOPFTIETT					
79	11	2926	SUSYSPGEI					
86	12	1651	SUTGRGRGSCIV					
100	14	2178	SMLTOPSHI					
100	14	2178	SMLTOPSHII					
86	12	2163	SOLCPPEPOV					
93	13	2209	SOLQASPL					
79	11	2209	SOLQASPLKA					
	11	2209	SOLQASPLKA					
	78							

UCV A02 Super Motif with Blinding Information

Conservancy	Flag	Position	Sequence	A'0201	A'0202	A'0203	A'0204	A'5802
93	13	56	SOPGRORPI					
86	12	1242	STKPAAYAA					
79	11	1242	STKPAJAYAA					
100	14	1784	STLPGNPA					
79	11	1784	STLPGNPAI					
79	11	2	STNPKQPKT					0.0007
86	12	1663	STWVLGGVY					
86	12	1663	STWVLGGVYL					
86	12	1663	STWVLGGVLA					
86	12	1299	STYGFELA					
100	14	1282	SVATLGFGA					
86	12	1455	SVDCNTCV					
86	12	1455	SVDCNTCVT					
86	12	995	TAACGDI					
86	12	1343	TGARLVL					
86	12	1343	TGARLVLV					
86	12	1343	TGARLVLVLA					
79	11	1343	TGARLVLVLAT					
79	11	2852	TARHPVNSWL					
79	11	2590	TIMAKNEV					
93	13	1268	TLGFAYM					
86	12	1268	TLGFAYMSKA					
79	11	1622	TLHGPTPL					
79	11	1622	TLHGPTPLL					0.0070
86	12	1611	TLFNLGGWV					
79	11	686	TULPLSTGL					
79	11	688	TULPLSTGLI					
79	11	1785	TLPGNPA					
86	12	125	TLTGFADL					
86	12	125	TLTGFAADM					
79	11	2871	TLWARMIL					
79	11	2871	TLWARMILM					
79	11	2671	TLWARMILMT					
86	12	1209	TMRSVF					
66	12	1464	TGT:CT:SL					
66	12	1464	TOTVFSLOPT					
79	11	2589	TTIMAKNEV					
79	11	885	TTLPALST					
79	11	685	TTLPALSTGL					
79	11	685	TTLPALSTGLI					
86	12	1208	TTMRSPVFT					
78	11	2738	TTSGGNIL					

LICV A01 Suber-Mohr with Binding Information

Conservancy	Freq.	Position	Sequence	A'0201	A'0202	A'0203	A'0206	A'6802
78	11	2739	TTSCGNTLT					
79	11	1597	TVCARAOA					
86	12	1466	TVFSLOOPT					
88	12	1466	TVFSLOFTFT					
100	14	1338	TVDOAET					
100	14	1336	TVDOAETA					
88	12	1336	TVDOETAGA					
100	14	1263	VAITLGFQA					
93	13	1283	VAITLGFAYM					
88	12	1230	VARLHAPT					
86	12	1440	VATDALMT					
86	12	1592	VAYQATVCA					
79	11	1592	VAYQATVCARA					
100	14	1420	VAYRGGLDV					
100	14	1420	VAYRGGLDVS					
86	12	1456	VIDCNTCV					
86	12	1456	VIDCNTCVT					
86	12	1456	VIDLTCGFA					
86	12	1422	VLAALAAYCL					
93	13	1671	VLCGYDA					
78	11	1521	VLCGYDAGCA					
100	14	1337	VLDQAEATA					
86	12	1337	VLDQAEATAGA					
86	12	157	VLEDGVNYA					
86	12	157	VLEDGVNYAT					
100	14	1258	VLNPSVAA					
100	14	1258	VLNPSVAAAT					
100	14	1258	VLTTSCGNT					
79	11	2737	VLTTSCGNTL					
78	11	2737	VLTTSCGNTLT					
79	11	1852	VLVDLAGYGA					
86	12	1666	VLVGGVLA					
88	12	1556	VLVGGVLA					
86	12	1866	VLVGGVLAAL					
86	12	1866	VLVGGVLAALA					
100	14	1256	VLVLPNSV					
100	14	1256	VLVLPNSVA					
100	14	1256	VLVLPNSVAA					
100	14	1256	VLVLPNSVAAAT					
79	11	2600	VOPEKSGRPA					

NCV A01 Super Motif with Binding Information

Conservancy	Freq.	Position	Sequence	A'0201			
				A'0202	A'0203	A'0206	A'6802
100	14	1916	VOVMANTU				
100	14	1918	VOVMANLIA				
100	14	1918	VOVMANLIAFA				
86	12	1463	VTOTVDFSL				
79	11	1138	VTRHADVI				
79	11	1138	VTRHADVIPV				
86	12	1661	VTSTWVLV				
86	12	1681	VTSTWVLVGV				
79	11	1439	VVATDALM				
79	11	1439	VVATDALMT				
79	11	1901	VVCALUINHV				
79	11	1698	VGVVCAAA				
79	11	1698	VVGVVICAAI				
79	11	1698	VVGVVCIAIL				
86	12	1660	VVTSTWVL				
86	12	1660	VVTSTWVLV				
86	12	1766	WAHKHMWNFI				
86	12	76	WICPGYTPWPL				
66	12	2873	WARMILMT				
79	11	2287	WAPOYNPPL				
100	14	1920	WHINLIAFA				
79	11	557	WNSTGFT				
88	12	1665	WWVNGGL				
66	12	1665	WWVSGVLA				
66	12	1865	WWVGGVLA				
66	12	1665	WWVGGVLAAL				
79	11	1249	YAAQGYKV				
79	11	1249	YAAQGYKVL				
79	11	1249	YAAQGYKVLV				
79	11	1249	YIPLVGAPL				
79	11	136	YLAGLSTL				
100	14	1779	YLGSSGGPL				
86	12	1185	YLGSSGGPL				
86	12	1185	YLGSSGGPL				
93	13	35	YLTTRHADVI				
79	11	2836	YLTTRHADTT				
86	12	1580	YLVAYOAT				
86	12	1590	YLVAYOATV				
86	12	1590	YLVAYOATVCA				
86	12	1138	YLVTRHADV				
79	11	1136	YLVTRHADVI				
93	13	1594	YQATVCARA				

HCV A02 Sanger Motif with Binding Information

Conservancy	Freq.	Position	Sequence	A'0201	A'0202	A'0203	A'0206	A'6802
79	1	1584	YQATVCARAQAA					
79	1	1106	YTNNDQDL					
79	1	1106	YTNNDQDLV					
86	12	276	YQDOLCESV	0.0018				
86	12	276	YQDOLCGSVFL					
93	13	637	YIGGVVERL	0.0008				
88	12	1939	YVPESDAA					
88	12	1939	YVPESDAAA					
86	12	1939	YVPESDAAVRV					
86	12	555						

Table IX
ICU Admissions and Mortality by Age Group and Form of Injury

Conservancy	Freq.	Position	Sequence	A'1101	A'3101	A'3301	A'8801
88	12	847	AACWYTRGER	0.0003	0.0140	0.0055	0.0018
79	11	147	AARALAHGVRA				
79	11	1187	ANVCTRGVAK				
79	11	2208	ASOKSAPSLK				
86	12	1265	ATLGFGAYMSK				
79	11	40	ATRKTSEF				
79	11	1186	AVCTRGVAK				
86	12	2841	CQLRQGVPRL	0.0260	0.0250	0.0011	0.0001
78	11	555	CTYMANSTGFK	0.7600	0.7500		
79	11	2598	CYOPEKGR	0.0008	0.0005		
79	11	2599	CYOPBKGRK	0.0011	0.0008		
100	14	1574	DAHLSOTK	0.0003	0.0005		
93	13	2617	DLGVRVCEK	0.0003	0.0002		
79	11	1143	DVPIVNR	0.0440	0.0440		
06	12	2245	EMGQHTTR				
86	12	2598	EVFCVOPK				
100	14	728	FLLDADH				
79	11	146	GAARALAHGVRA				
100	14	1918	GANGWMMR				
79	11	3037	GYTLPNH				
79	11	1004	GLPVSAFR				
85	12	1131	GSSDLYVTR				
86	12	1863	GVAGALVAFK	0.3900	1.4000	0.0055	0.0088
79	11	3036	GYVYLPNQ	0.0014	0.0140	0.1600	0.0007
79	11	45	GYFRATKTSER				
79	11	1900	GYVCAAILA				
79	11	1900	GVVCAAILRA				
63	13	33	GYTLPRIN				
93	13	33	GYVLPFRGPR				
79	11	1141	HADYFVRA				
78	11	1141	HADYFVRA				
79	11	1141	HADYFVRA				
100	14	1234	HAPTGSCK				
93	13	1234	HAPTGSCKTK				
100	14	1572	HIDAHRSOTK				
86	12	1232	HJHAPTGSCK	0.5900	0.0024	0.0005	0.0028
100	14	1395	HJIFCHSK	0.0250	0.0006	0.0003	0.0010
100	14	1395	HJIFOISKK	0.0260	0.0002	0.0009	0.0001
79	11	2920	HSYSQGENR				
79	11	222	HPGCQPCVRA	0.0004	0.0012		
86	12	2250	ITIVSEENK	0.0150	0.0079	0.0007	0.0002
86	12	1288	ITVSYTK				
78	11	2813	MFPOLGIV	0.0036	0.0044		
93	13	30	MGGVYLLPRA	0.0008	0.0056		
93	13	30	MGGVYLLPRA				
86	12	2944	KLQVPPAR				
86	12	10	KTKRANTRA				
86	12	10	KTKRANTRA				
93	13	51	KTSENSOPN	0.0110	0.0100		
93	12	51	KTSENSOPN	0.0840	0.2700	0.0160	0.0850
86	12	1128	LAEGDFK				

ICV A01 Stutter Motif (With Binding Information)

Conservancy	Freq.	Position	Sequence	A'0301	A'1101	A'3101	A'3301	A'8801
86	12	2235	LEANLWRA	0.0008	0.0005	0.0018	0.0068	0.0008
100	14	1398	LIFCHSKK	0.5400	0.1800	0.0071	0.0012	0.0240
100	14	1598	LIFCHSKK	0.0003	0.0001			
79	11	2612	UVFFPOLGVA					
100	14	726	LFLLLADAR					
93	13	38	UPPROPA					
86	12	87	USPRCSR					
79	11	1591	LVAYQATVCAR					
79	11	1	MSTNPKPQR					
79	11	1	MSTNPKPQRK					
86	12	2248	NTRVESENK					
79	11	14	MTRNPDKK					
78	11	1295	PTYSTGK					
79	11	2667	PNIGFSYDTA					
83	13	614	PSAVVGTIDR					
79	11	1607	PSWDDMMK					
86	12	109	PTOPRDSR					
93	13	1230	PTGSKSTK					
93	13	616	PVYVGTTDR					
86	12	1340	QAEFTAGAR					
93	13	29	CNGGKYLPR					
86	12	288	QLFTESPR					
78	11	289	QLFTESPR					
78	11	2210	OLAPSILK					
79	11	1186	RAAVCTGIVAK					
100	14	149	TAUAIQVR					
79	11	47	PATIKTSEK					
79	11	43	PLCVRATH					
79	11	43	PLQVATRKK					
70	11	1923	RLIAFASR	0.8400	0.0280	0.0420	0.0004	0.0001
100	14	2811	RLVFFPOLGVA					
79	11	635	TMNGCVERIN					
100	14	635	TMNGCVERIN					
93	13	55	TSQFTGPR					
79	11	2207	SASCLASPLK					
86	12	1132	SSDOLVTR					
79	11	2	STRPKPQR					
79	11	2	STRPKPQRK					
79	11	2	STRPKPQRK					
86	12	1268	TLGFOAYMSK					
79	11	1622	TLGFTPLYR					
93	13	62	TSERSOPH					
86	12	52	TSNSOPHGR					
90	12	62	TSNSOPHGR					
86	12	1050	TSFLGRK					
86	12	1684	VAGALYAFK					
79	11	1582	VAYGATVCAR					
93	12	1337	VLDDAETAGAR					
79	11	1138	VTRHADVIPVR					
79	11	1901	WVCAAILR					
79	11	1901	WVQVCAAILR					
79	11	1080	WVQVCAAILR					
93	13	517	WVGTTR					

HCV_A01 Super Model (With Bimutate Information)

Conservancy	Freq.	Position	Sequence	A'0301	A'1101	A'3101	A'3301	A'6801
86	12	93	WAGWMLSPR					
86	12	96	WILSPRSH	0.0008	0.0005			
100	14	1920	WMNRLIAFASR					
79	11	557	WMNSTGFTK	0.0530	0.0010	0.0014	0.0420	0.0058
93	13	35	YLPRRGPR	0.0054	0.0005			
79	11	2930	YSPGENR					
100	14	637	YGGVEPR					
86	12	1939	YWPESDAAV	0.0003	0.0001			
			112					

Table X
HCV A24 Super Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A ²⁴⁰¹
ALSPGAL	1890	6	9.3	1.3	
ALAHGVRL	150	9	1.4	1.4	
ALSTGLIHL	669	9	1.2	1.2	
ALVVGVVCAAI	1896	11	7.9	1.1	
ATGLNPGCSF	165	10	1.3	1.3	
ATLGFGAY	1265	6	1.2	1.2	
ATLGFGAYM	1265	9	1.3	1.3	
AVAYYRGL	1419	8	1.4	1.4	
AVQWMNRFL	1917	6	1.4	1.4	
AVQWMNRFLU	1917	11	1.4	1.4	
AVQWMNRFLAF	1917	11	1.4	1.4	
AVDMMMMW	319	6	1.2	1.2	
AYAAQGYKVL	1248	10	1.1	1.1	
AYFAGLDVSVI	1421	11	1.4	1.4	
CIRKLGVPPPL	2941	10	1.2	1.2	
CLVMMLLI	739	6	1.2	1.2	
CTCGSSSL	1128	8	1.1	1.1	
CTCGSSSDLY	1128	9	1.1	1.1	
CTCGSSSDYL	1128	10	1.1	1.1	
CTFGVAKAVDF	1190	11	1.1	1.1	
CTWMNSTGF	555	9	1.1	1.1	
CVTOTVDF	1462	8	1.2	1.2	
CVTQVDFSL	1462	10	1.2	1.2	
CVDAGCAGW	1525	8	1.1	1.1	
CVDAGCAWY	1525	9	1.1	1.1	
CVDAGCAWYEL	1525	11	1.1	1.1	
DFSLOPTFT	1468	8	1.4	1.4	
DFSLOPTFTI	1468	10	1.4	1.4	
DIGSVRL	279	8	1.2	1.2	
DLEWVTSW	1657	9	1.2	1.2	
DLEWVTSWVL	1657	11	1.2	1.2	
DLGIVRCER	2617	10	1.3	1.3	
DLMGYIPL	132	8	1.1	1.1	
DLYNLLPAL	1683	9	1.1	1.1	
DVNLNLPAI	1883	10	1.1	1.1	
DYKPGGGGQ	21	10	1.2	1.2	
DYVYALWY	616	8	1.4	1.4	
DYVYALWYH	616	9	1.2	1.2	
DTAACGDI	994	9	1.2	1.2	
DTAACGDI	994	10	1.2	1.2	
DTLTGCFADL	124	11	1.2	1.2	
DTLTGCFADL	124	12	1.2	1.2	
DWKPQGGGQ	21	10	1.2	1.2	
EPFYGLWY	1377	9	1.3	1.3	
EPFYGLWYH	1377	10	1.2	1.2	
ETAGARLVL	1342	9	1.2	1.2	
ETMMSPVF	1207	12	1.2	1.2	
EVTSVWL	1659	9	1.2	1.2	

UCCV Δ 24 Super Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A' 2401
FGSGGNL	1773	8	14	100	
FGSQYLAGL	1773	11	14	100	
FLALLSCL	177	9	12	86	
FTEAMTRY	2782	6	14	100	
FTGLTHDAHF	1667	11	13	93	
FTTLPALSTGL	684	11	11	79	6.9000
FWAQHANIF	1765	9	12	86	
FWAKHHWNIF	1765	10	12	86	
GFADUMGJY	129	8	13	93	
GFADUMGJYI	129	9	13	93	
GFADUMGJYPL	129	11	11	79	
GFSYDTRCF	2689	9	11	79	
GIOYLAGL	1776	8	14	100	
GIOYLAGLSTL	1776	11	14	100	
GLPVQODHL	1652	0	13	93	
GLPVQODHLF	1552	11	12	86	
GLSAFSLJSY	2921	10	11	79	0.0001
GLSTLQGRNF	1782	11	11	79	
GLTHDAHF	1589	9	13	93	
GLTHDAHF-L	1589	10	13	93	
GTFPINAY	2063	8	11	79	
GVAGSALVAF	1863	9	12	86	
GVAKAVDF	1193	8	11	79	
GVLAALAY	1670	9	12	86	
QVLAALAAVCL	1670	11	12	86	
QVNTATGQL	161	8	11	79	
GVRICEM	2618	8	14	100	
GVVCEKMAI	2619	10	14	100	
GVVCEKMAIY	2619	11	14	100	
GVRLEDQWY	154	11	12	86	
GWVCAAIL	1900	8	11	79	
GWILLAP	1027	8	11	79	
QWRLAPIAY	1027	11	11	79	
QYGAQYAQAL	1859	10	12	86	0.0003
QYIPVGAPL	135	10	11	79	0.0057
QYFRGRASCVL	2728	11	12	86	
HJHQHNDWY	698	11	11	79	
HLPNEQGM	1719	9	11	79	
HMWNFSGL	1769	9	13	93	
HMWNFSGQY	1789	11	13	93	
HTPNWSWL	2855	8	12	86	
HTPNWSWLII	2855	11	12	86	
HVGCEBQW	1910	10	12	78	
IFLLALLSCL	176	8	12	86	
ILGGWVAAQ	1816	10	12	86	0.0026

HCY A24 Super Motif With Blinding Information

Sequence	Position	No. of Amino Acids	Frequency	Conservancy (%)	A ²⁴⁰¹
ILGIGIVL	1331	8	12	86	
IMAKNEVF	2591	8	12	86	
ITYSTQKF	1296	9	12	86	
ITYSTQKFL	1296	10	11	79	
NDVQMLY	701	8	12	86	
NGANML	30	8	13	93	
KPGQQQI	23	8	13	93	
KVIDLTQDF	121	10	12	86	
LPNLGSW	1813	8	12	86	
LIEANLILW	2235	8	12	86	
LINTNGSW	414	8	11	79	
LLAISCL	170	0	12	86	
LLAPITAY	1030	8	14	00	
LPNLGGW	1812	8	12	86	
LIPALSPGAL	1887	11	13	83	
LPRGPGPL	36	9	13	83	
LSPRGSPSPSN	97	11	11	79	
LYMPDGEN	2240	11	12	06	
LTCGPADL	126	8	12	86	
LTCQFADLM	126	9	12	86	
LTCQFADWGY	126	11	12	86	
LTHIDANIF	1570	8	13	93	
LTHIDANFL	1570	9	13	93	
LTSMALDPSHI	2176	11	13	93	
LTSSCANTL	2738	8	11	79	
LYDIAQY	1853	8	11	79	
LYGVALAL	1687	9	12	86	
LVNPQVWATL	1257	11	14	100	
LVNLIPAI	1804	8	11	79	
LVNLIPAI	1884	8	11	79	
LYTRHADI	1137	9	11	79	
LVGVWCAAI	1897	10	11	79	
LVGVWCAIL	1897	11	11	79	
LWARMILM	2872	8	12	86	
LWARMILMTHF	2872	11	12	86	
LWRCERGGN	2241	10	12	86	
LYLYTRHADI	1135	11	11	79	
MILMTHF	2878	8	12	86	
MLTDFSHI	2179	8	14	100	
MNNFISGI	1770	8	14	100	
MNNFISQY	1770	10	14	100	
MNNFISQY	1770	11	14	100	
MNGGVEFL	638	10	13	93	0.0270
NFSQY	1772	8	14	100	
NFSQY	1772	9	14	100	0.0170

ICV A24 Super Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A ²⁴ 01
NIIGGWWAQL	1815	11	12	86	
NIRTQVPTI	1282	9	11	79	
NIVDQVY	700	8	12	86	
NIVDQVLY	700	9	12	86	
NLGKVIDTL	116	9	12	86	0.0001
NLWDRDEM	2235	8	12	86	
NLPGCSFESI	168	9	13	93	
NLPGCSFSIF	168	10	13	93	
NLPGCSFPL	168	11	13	93	
NTCVTQTDIF	1480	10	12	86	
NTNGSWHI	416	8	13	93	
NTNRDQVKF	14	11	11	79	
NTDQDQVGN	1108	9	11	79	
NWFQCTWM	561	8	12	86	
PTIYSTYKWF	1295	10	11	79	
PTIYSTYKFL	1295	11	11	79	
PLEGERQDPQL	2403	11	13	93	
PLGGAALARL	143	9	11	79	
PMGSDYDRCF	2867	11	11	79	
PTDPRHRSNL	109	11	12	86	
PTLHGPPPL	1621	9	11	79	
PTLHGPTPLL	1821	10	11	79	
PTLHGPTPLY	1821	11	11	79	
PTLWARMI	2870	8	11	79	
PTLWARMILM	2870	9	11	79	
PTPLYRL	2870	10	11	79	
PVCOOHIEF	1626	8	14	100	
PVCOOLIEFW	1554	9	12	86	
PVNSMLGNI	2867	10	12	86	
PVNSMLGNI	2857	9	14	100	
PVNSWLGRNIA	2557	10	14	100	
PVNHGCP	2318	8	11	79	
QFKKAKGL	1732	9	12	86	
QFKKAKGL	1732	10	12	86	
QVGGVYI	29	8	13	93	
QVGGVYI	29	9	13	93	
QVDFSLDPTF	1465	11	12	86	
QWANRLLAF	1919	9	14	100	
QYLAGLSTL	1778	9	14	100	0.0480
QYSGQREF	2847	10	11	79	0.0180
QYSGQREFL	2847	11	11	78	
FLHGLSAF	2918	8	12	86	
FLHGLSAF	2918	10	11	79	
FLHGLSAF	2611	8	11	79	

HCV_A24_Super Motif With Binding Information

Sequence	Position	Peptide No.	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'2401
RLLAPITAY	1029		9	12	66	
RMANDMM	317		8	12	66	
RMADWMMNNW	317		10	12	66	
RMLMTIF	2875		8	12	66	
RMLMTIFF	2875		9	12	66	
PRVYQGVERL	835		11	12	86	
RVCEKMAAL	2621		8	13	91	
RVCEKMAALY	2821		9	14	100	
RMLEDGNY	156		9	14	100	
SFSFILAL	173		9	12	86	
SFSFILALL	173		9	14	100	
SIFLALL	175		8	14	100	
SIFLALLSCL	175		11	12	86	
SLOPFTFI	1470		8	14	100	
SUHSYSPGIEI	2928		10	11	100	
SMLTOPSPH	2178		0	14	100	
STKVPAAV	1242		0	12	66	
STLPGNPAN	1784		9	11	79	
STWMLVGGVL	1883		10	12	86	
SVAATLGF	1262		8	14	100	
SVAATLGFQAY	1262		11	14	100	
SYQDMMWCL	1808		9	11	79	
SYLGNINM	2860		0	12	66	
SYLKGSQGQL	1164		11	12	86	
TIMAKNEVF	2590		9	11	79	
TLOFGAYM	1285		8	13	91	
TMQPTFL	1622		8	11	79	
THGPIPALL	1622		9	11	79	
TUQPTPLY	1622		10	11	79	
TLPNLGAW	1811		10	12	66	
TLPALSTGL	666		9	11	79	
TLPALSTGL	666		10	11	79	
TLPGNPAI	1786		8	11	79	
TLCGFAQL	125		9	12	86	
TLCGFAADM	125		10	12	86	
TLMARMIL	2871		8	11	79	
TLMARMIL	2871		9	11	79	
TIMAKNEVF	2589		10	11	79	
TLPALSTGL	885		10	11	79	
TLPALSTGL	885		11	11	79	
TMRSQPF	1208		8	12	86	
TTSCGNTL	2739		8	11	79	
TVDSLQPTF	1466		10	12	86	
TWANSTGF	556		8	11	79	
TWANSTGF	1664		9	12	86	

HCV A24 Super Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*2401
TYSTYGRKF	1287	8	13	93	
TYSTYGRKF	1287	9	12	86	
VFTGLTHI	1586	8	13	93	
YDITLTCGF	122	9	12	86	
VLAALAY	1871	8	12	86	
VLAALAAYCL	1671	10	12	86	0.0070
YLEDGVNY	167	8	12	86	
VINPSVATL	1258	10	14	100	
VLTTSGCNTL	2737	10	11	79	
VLVBDLAGY	1852	9	11	79	
VLVGGVAAAL	1868	10	12	86	
WGGSSYGF	2839	8	11	79	
WGGSSYGFY	2639	10	11	79	
VIQNIVDFSL	1463	9	12	86	
VTRHADVI	1138	8	11	79	
WVATDALM	1439	8	11	79	
WGIVCAAI	1698	9	11	79	
WVAVCAAIL	1890	10	11	79	
WTSTWVL	1660	8	12	86	
YVLPFRGFL	34	11	13	83	0.0016
WMAARHLAF	1920	8	14	100	
WWLVCSCV	1685	8	12	86	
WWLVSGVLAAL	1665	11	12	86	
YIPLVSGFL	158	9	11	79	
YLAGLSTL	1779	8	14	100	
YLGSSGGFL	1185	10	12	86	
YUKSSGGFL	1185	11	12	86	
YLPRPRGPFL	36	10	13	93	0.0001
YLVTFRADVI	1138	10	11	79	
YTNVDOOL	1108	0	11	79	
YTNVDOOLGVW	1108	11	11	79	
YVGLOGSVF	276	10	12	86	
YWDLGGSVFL	276	11	12	86	
YVGVBVR	637	9	13	93	
YRGGLDVSFV	1422	10	14	100	
		3			

Table XI
UCY.DAT Super Matrix (with Bimutate Information)

Conservancy	Freq.	Position	Sequence	B'0702	B'3501	B'5101	B'5301	B'5401
86	12	1604	APPSSWDDMM	0.00026	0.0002	0.0002	0.0001	0.0002
79	11	1604	APPSSWDDMM	0.0001	0.0001	0.0002	0.0006	0.0003
93	13	1235	APTGSGKSKTV	0.0001				
79	11	2869	APTLWARM	0.4300	0.0001	0.0012	-0.0002	0.0023
79	11	2869	APTLWARMIL	0.0160	0.0002	0.0012	0.0001	0.0002
79	11	2869	APTLWARMILM	0.0160	0.0001	0.0010	0.0001	0.0003
79	11	2869	DPOLSGGSW	0.0001	0.0001	-0.0003	-0.0002	0.0033
79	11	2410	DPRFASRNL	0.0001	0.0002	0.0002	0.0005	0.0002
86	12	111	FPDQGIVR	0.0170	0.0002	0.0001	0.0001	0.0002
79	11	2815	FRGGGCV	0.0001				
100	14	24	FPGGCGNGV	0.0001				
86	12	1912	GPGEGAMDN	0.0001	0.0002	0.0001	0.0001	0.0002
86	12	1912	GPGEGAWMM	0.0001	0.0001	0.0002	0.0001	0.0003
93	13	41	GPTLGIVRA	0.0001				
100	14	1625	GPFLPLVYL	0.0024	0.0002	0.0002	0.0001	0.0002
93	13	607	GPTFLYRLGAG	0.0001				
93	13	1370	GPYCFTPSPV	0.0001				
79	11	137	IPFYGKAI	0.0120				
86	12	2608	IPVYGAFL	0.4400	0.0032	0.0700	0.0003	0.0035
79	11	2608	KPARLIVF	0.0150	0.0002	0.0017	-0.0002	0.0008
79	11	2608	KPABLVFPDL	0.0003				
78	11	1620	KPTLHGPPFL	1.4150	0.0001	0.0002	0.0001	0.0003
79	11	1620	KPTLHGPTPL	0.0021				
93	13	1088	LPAILSPGA	0.0001	0.0001	0.0001	0.0002	0.0400
93	13	1088	LPAILSPQAL	0.0053	0.0001	0.0036	0.0001	0.2100
60	12	1088	LPAILSPGALV	0.0003				
100	14	607	LPALSTGL	0.0020				
86	12	607	LPALSTGLI	0.0350	0.0002	2.0000	0.0082	0.0005
86	12	607	LPALSTGLHL	0.0011				
86	12	2165	LPCEPEPDV	0.0001	0.0002	0.0001	0.0001	0.0002
93	13	109	LPGESFSL	0.0110	0.0360	0.0560	0.0150	0.0018
93	13	169	LPGCSFSF	0.1950	0.0790	0.0550	0.0013	0.0015
93	13	169	LPGCSFSFL	0.0022	0.0008	0.0100	0.0140	0.0012
93	13	169	UPHQQLP	0.0007				
93	13	37	UPRGPFLGV	6.5000	0.0001	0.0180	-0.0002	0.0020
93	13	1553	UPRGPFLH	0.1900	0.0001	0.0009	0.0001	0.0025
93	13	1553	UPVQDHLF	0.0005				
86	12	1553	UPVQDHLFW	0.0001				
86	12	1720	LPTEGAM	0.0130	0.0001	0.0040	-0.0002	0.0013
100	14	1260	NPSVATL	0.0011				
100	14	1260	NPSVATLGF	0.0001	0.0001	0.0002	0.0001	0.0003
86	12	1605	PPPSWDDOM	0.0003				
79	11	1605	PPPSWDDMW	0.0001				
79	11	1608	PPPSWDDMM	0.0002				
79	11	1608	PPPSWDDMMC	0.0001				
79	11	2317	PPVYHGPPL	0.0140	0.0001	0.0001	0.0001	-0.0002
79	11	2801	CPKCGTRPA	0.0111	0.0001	0.0002	0.0190	0.0190
79	11	2809	OPETQEL	0.0002				
70	11	2808	OPETQEL	0.0001	0.0002	0.0001	0.0001	0.0002
86	12	78	OPGTPWPL	0.0006				

JUCY.D07 Super Multi (with Binding Information)

Conservancy	Freq.	Position	Sequence	B*702	B*3501	B*5101	B*5301	B*5401
86	12	76	OPGYMPWLY	0.0001	0.0002	0.0001	0.0002	0.0002
93	13	57	OPGRGPQPI	0.2300	0.0002	0.0001	0.0001	0.0002
79	11	2289	RFDYNFPL	0.0050	0.0001	0.0002	0.1200	0.0002
93	13	1893	SPGALVGV	0.0001	0.0002	0.0001	0.0016	0.0003
79	11	1883	SPGALVGVV	0.0130	0.0001	0.0001	0.0001	0.0003
79	11	2931	SPGEINRV	0.0007	0.0003	0.0001	0.0002	0.0037
79	11	2931	SPGENRVA	0.0003	0.0001	0.0001	0.0002	0.0002
79	11	2849	SPGREF	0.0027	0.0002	0.0002	0.0001	0.0002
79	11	2849	SPGREFEL	0.1200	0.0002	0.0002	0.0001	0.0002
79	11	99	SPGSRPSW	0.3800	0.0002	0.0005	0.0001	0.0002
86	12	1935	SPTHYVPESDA	0.0001	0.0028	0.0001	0.0001	0.0003
86	12	1975	TPCSGSW	0.0005	0.0001	0.0002	0.0001	0.0003
79	11	1126	TPCTCGSSOL	0.0001	0.0001	0.0002	0.0001	0.0003
79	11	1126	TPCTCGSSOLY	0.0001	0.0001	0.0002	0.0001	0.0003
86	12	223	TPGCYFCV	0.0001	0.0001	0.0002	0.0001	0.0003
93	13	1650	TPGLPVQDCHL	0.0001	0.0003	0.0001	0.0002	0.2300
93	13	1627	TPLYRIGA	0.0003	0.0001	0.0001	0.0001	0.0110
85	13	1827	TPLYRIGAV	0.0120	0.0001	0.0008	0.0001	0.0003
86	12	2056	TPNSWLGNI	0.0001	0.0001	0.0053	0.0008	0.0003
86	12	2856	TPNSWLGNI	0.0001	0.0022	0.0001	0.0010	0.0003
86	12	1940	VPESDAAARV	0.0001	0.0001	0.0001	0.0001	0.0003
86	12	1940	WPLLLLL	0.0021	0.0001	0.0010	0.0001	0.0003
86	12	799	YPYRLWHY	0.0001	0.0001	0.0001	0.0001	0.0003
100	14	616	76					

Table XII
IICV_B27 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
AKHMINIF	1767	8	12	86
AKNEVFCV	2593	8	12	86
ARALAHGV	148	8	14	100
DRSESPV	663	8	8	11
EKGGRGPA	2603	6	11	79
EKMALYDV	2624	8	12	79
FKGKALGL	1733	8	12	86
GHRMAMOM	315	8	12	86
GKSTKVPK	1240	8	13	93
GRKPARU	2806	8	12	86
HRMAMOMMA	316	8	13	79
KCGGRHLI	1390	8	11	93
IRATGVRT	1283	8	11	79
KKCDELAA	1403	8	14	100
KRKGDELAA	1402	8	14	100
LH-GPTPLI	1623	8	11	70
LH-GNNDV	697	8	12	86
LIDLAVAY	969	8	11	78
NHVSFPHY	1932	8	12	86
PGRGRGP	56	8	13	93
PGRSRPSW	100	8	11	79
PRRSRPNL	112	8	12	86
RHAQDIPV	1140	8	11	79
RHTPNWSW	2854	8	12	86
RKLGVPPL	2943	8	12	86
RKPARLIV	2607	8	11	79
RICRASQV	2730	8	12	89
RFGFILGV	39	8	13	93
RPGQDMKF	17	8	12	86
SKKQDEL	1401	8	14	100
SPNLGAVI	118	8	12	86
THIDAHFL	1571	8	13	93
TKLKLTPI	2985	8	12	86
TKVPAAYA	1243	8	12	86
TRCFDSTV	2674	8	14	100
TRGVAKAV	1181	8	11	79
VRVCEKMA	2620	8	14	100
VRMLEDGV	155	8	13	93
YRGLOVSV	1423	8	14	100
AFHHTPVASW	2853	8	11	79
ARLIVFPDL	2810	9	11	79
ARLVVLTAA	1346	9	11	79
ARMILMTHF	2874	9	12	86
ARPDYNPPPL	2298	9	79	79
DRSESPV	663	9	11	79

HCV R27 Super Motif

Sequence	Position	Peptide No.	No. of Amino Acids	Sequence Frequency	Conservancy (%)
EKMLYDVV	2624		9	12	8.6
FKKAKLGLL	1733		9	12	6.6
GHRMADM	315		8	13	9.3
GKSTKVPAA	1240		9	12	8.6
GRKPARLV	2608		9	11	7.9
HRKAWDMM	316		9	12	8.6
IGGGHLIF	1390		9	11	7.9
KKKCDELAA	1402		9	14	10.0
UHGLSAFSL	2919		9	11	7.9
UHGPTPLLY	1623		9	11	7.9
LHYSPGEI	2927		9	11	7.9
LKGSGGGPL	1166		9	12	8.6
LRLKGVPPL	2942		9	12	8.6
NHVSPTHIV	1932		9	12	8.6
NRPPDQW	16		9	11	7.0
PRGRPLGLV	38		9	13	8.3
RHTPVNSW	2854		9	12	0.6
RHVGPGEGA	1909		9	11	7.9
RKPKARLIVF	2607		9	11	7.9
RKRCRSGVL	2730		9	12	6.6
RRSRNALGV	114		9	12	6.6
SKKKCDELA	1401		9	14	10.0
THYVPESDA	1937		9	12	6.6
TKVPAAYAA	1243		8	11	7.9
TRHADIVRV	1139		9	11	7.9
TRVSEENKV	2251		9	12	8.6
VKPGGGGAI	22		9	13	9.3
VVICCEKML	2620		9	14	10.0
WRLLAPITA	1028		9	11	7.9
WRCBAGGN	2242		9	12	8.6
YRGDLSVI	1423		8	14	10.0
YPRCRASCV	2129		9	13	9.3
ARALAHGVRY	148		10	14	10.0
ARADQAPPNSV	1600		10	11	7.9
ARHTRPNSWL	2853		10	11	7.9
ARMILATHFF	2874		10	12	8.6
CRSKKCDEL	1399		10	14	10.0
DRDRSEL SPL	861		10	11	7.9
DRSELSPLL	663		10	11	7.9
EVGGBVRPL	2664		10	11	7.9
FRAAVICTRGV	1185		10	12	8.6
GHRMADM	316		10	12	8.6
GKSTKVPAA	1240		10	12	8.6
GRKPARLV	2606		10	11	8.6
KHAWNIFSGI	1768		10	13	9.3

HCV_B27 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
KKCDDELAALK	1403	10	12	86
LHONIVDVOY	697	10	11	79
LGSSSGGFL	1168	10	12	86
OKALGLQATA	1735	10	12	86
PRVGRGEGAV	1809	10	11	79
PRSPALVRA	39	13	13	93
PRHNGPGEA	1908	10	11	79
PRPSPLGKV	113	10	12	86
RASFLALGKV	114	10	12	86
SFGYGAQDVK	2552	10	12	86
SKKKCDELA	1401	10	14	100
THVPESEDA	1937	10	12	86
TRGVAKAVDF	1191	10	11	79
TRVSEENKVA	2251	10	12	86
VKPGGGGCV	22	10	13	93
VVICERKMLY	2620	10	14	100
VRMLEDWNN	155	10	12	86
WRLLAPITAY	1028	10	11	79
YKVLNLPSV	1254	10	14	100
YRCRASGV	2726	10	12	86
AIGVRALEGV	152	11	13	83
AKHMMWNFSGI	1767	11	12	86
ARALAHGVRL	148	11	14	100
ARLIVFFPDGLV	2810	11	11	79
CHSKRKCDDEA	1399	11	14	100
DPDSESELSPL	661	11	11	79
EKGGRKPARL	2603	11	11	79
FRAAVCTRGVA	1105	11	11	79
GKSTKVPAAYA	1240	12	12	86
GKVIDLTGCF	120	11	11	79
HMAYWOMMMW	316	12	12	86
KKCDDELAALK	1402	12	12	86
KANTNRPPODV	16	12	12	86
LHGIFTPLYRL	1623	12	12	86
LHONIVDVOY	697	11	11	79
LKPTLHGPTPL	1619	11	11	79
LRHNGPGEA	1907	11	11	79
PRGPRGIVRA	38	13	13	93
PRPSPLGKV	112	11	12	86
PRVGRGEGAV	1908	11	11	79
PRVSGAGKV	113	11	12	86
SRGNHNSPHTY	1829	11	12	86
SRNLGVDTL	116	12	12	86
THVPESEDA	1937	11	12	86
VRVLEDGVNYA	155	12	12	86

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
YKVLVLNPSVA	1254	11	14	100

HCV R27 Super Motif

HCV_B58 Super Motif

Table XIII

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
AAILRRIV	1804	8	13	93
AALAYCL	1673	8	12	88
AADGYKVL	1250	8	11	79
AATLGFGA	1264	8	14	100
AAVCTRGV	1187	8	12	88
ASLMFTA	1783	8	11	79
ASSFASQL	2204	8	14	100
ATLGFAY	1265	8	14	100
CSFSFL	172	8	14	100
CGGGAYDI	1310	8	12	86
CSSNVSA	2619	8	14	100
CTCGSSQL	1128	8	11	79
CTRGVAKA	1180	8	11	79
DTAACGDI	994	8	12	86
DTLTCGFA	124	8	12	86
EAALENLV	750	8	11	79
EAMTRYSK	2794	8	14	100
ESDAAARV	1942	8	12	86
ETAGQRLV	1342	8	12	86
ETTMARSPV	1207	8	12	88
FAOLQYI	130	8	13	93
FASRGNHIV	1927	8	14	100
FSIFLLAL	174	8	14	100
FSYDTRCF	2670	8	11	79
FTEAMTRY	2792	8	14	100
FTPSPVN	512	8	13	93
GAGVAGAL	1861	8	12	86
GAMIGVLA	350	8	12	86
GALVGVW	1895	8	11	79
GARLVLA	1345	8	12	86
GSGKSTRV	1238	8	13	93
GSSDQYLV	1131	8	12	86
GSSGGFL	1168	8	12	86
GSSYGFQY	2841	8	11	79
GTFPINAY	2083	8	11	79
HSYSPGEI	2926	8	12	86
HTPNNSWL	2855	8	12	86
ISGIOYLA	1774	8	14	100
ITCSSSNV	2816	8	14	100
ITWGAQTA	989	8	12	86
KSTKVPAK	1241	8	12	86
LAGYGAQV	1857	8	11	79
LAHGIVRL	151	8	14	100
LAVAEPV	972	8	11	79
LSAPSLKA	2211	6	11	79

ICV B58 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
LSPGALVV	1892	8	13	93
LSTGGLHL	690	8	12	86
LTCGFADL	126	8	12	86
LTHIDAHF	1570	8	13	93
MSADLEVV	1654	8	11	79
NSWLGNII	2859	8	14	100
NTCVTQTV	1460	8	12	86
NTNGSWH	416	8	13	93
PAILSPGA	1889	8	13	93
PALSTGIL	688	8	12	86
PTLWARM	2870	8	11	79
PTPLLYRL	1626	8	14	100
QATVCARA	1595	8	13	93
RARPRWFM	3019	8	14	100
ASELSPILL	664	8	11	79
RSRNGLKV	115	8	12	86
SAPSUHSY	2923	8	11	79
SSASQLSA	2205	8	14	100
STKVPAAV	1242	9	12	86
STLPGNPA	1784	8	14	100
STLPOQAVM	2633	8	12	86
STYGKFLA	1299	8	12	86
TAACGDI	895	9	12	86
TAGAFLVY	1343	8	12	86
TTMRSVPF	1208	8	12	86
TTSGQNTL	2739	8	11	79
VAGALVAF	1864	8	12	86
VTRHADVI	1138	8	11	79
VTSTWMLV	1661	8	12	86
WAKHAWNPF	1766	9	12	86
WAKYLVNM	368	8	14	100
WAQGYPWN	76	8	12	86
YAAQGYKV	1249	8	11	79
YSEPLDL	2905	8	11	78
YSTYGKFL	1286	8	12	86
YTNAQGQL	1106	8	11	79
AAKLODCTM	2758	9	16	114
AAQGYKV	1250	9	11	78
AARALAHGV	147	9	11	79
ATLGFQAV	1264	9	14	100
AAVCTRGVA	1187	9	11	79
ASGLSAFSL	2208	9	13	83
ATLGFGAYN	1265	26	26	186
ATVCAHQA	1586	9	11	79
CAAMLRBH	1903	13	13	93

HCV 1558 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
CAWYELTPA	1630	9	11	79
CSFSFLLA	172	9	14	100
CGGGAYDN	1310	9	12	86
CTCOSSDLY	1128	9	11	79
CTRGVYAKAV	1190	9	11	79
CTWMNSTGF	555	3	11	79
DAGCAWYEL	1527	11	11	79
DTAACGDI	994	9	12	86
DTRCFDSTV	2673	9	13	93
ETASBARLW	1342	9	12	86
ETTMFISPVF	1207	9	12	86
FSIELLALL	174	9	14	100
FSIOPFTI	1469	9	14	100
FTGLTHIDA	1567	9	13	93
GAGVAGALV	1861	9	12	86
GALVAFKIM	1866	9	12	86
GALVAFKYM	1866	9	14	100
GAVQHANR	1816	9	14	100
HSKKKCDEL	1400	9	14	100
HTTPCCVPCV	222	9	11	79
ITWGADTA	989	9	12	86
ITSYTGKF	1296	9	12	86
KALGLQTA	1736	9	12	86
KSTKVPAAV	1241	9	12	86
LAALAAAYCL	1672	9	12	86
LAEOFKOKA	1729	9	12	86
LAGLAYISM	356	9	14	100
LAQYGAGVA	1857	9	11	79
LSAFSLTISY	2922	9	11	79
LSLPGNPA	1783	9	14	100
LTCGFADLM	126	9	24	171
LTDPFSHTA	2180	9	14	100
LTERDKHNV	1052	9	12	86
LTHIDARFL	1570	9	13	93
LTSSCGNTL	2738	9	11	79
MAKNEVFCV	2592	9	12	86
MAMDMMMMW	318	9	12	86
NAVAYYRGL	1418	9	13	93
NSLLH-FHM	2481	9	14	100
NSWLGNIM	2858	9	24	171
NTNRPPODV	14	9	12	86
PALSIGAL	1889	9	13	93
PSVAATLGF	1261	9	14	100
PTLWGPPTPL	1621	9	11	79
PTLWARMIL	2870	9	11	79

HCV 1558 Surfer Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
QAETGAGR	1340	9	12	86
RAAVCTGRV	1186	8	12	86
RALANGVRV	148	9	14	100
RAQAPPSSW	1601	9	11	79
RAYANDREW	811	9	16	114
RSELSPLL	664	9	11	79
RSENLGKIV	115	9	12	86
SSSASQSLA	2205	9	100	100
STKIPAYA	1242	9	14	12
STLPGNPAI	1784	9	11	79
STYVLVGGV	1663	9	12	88
TAGARLVLV	1343	9	12	86
TSCSSINSV	2817	9	14	100
TTIMAKNEY	2389	9	11	79
VAATLGFGA	1263	9	14	100
VAGGATVOM	933	9	14	100
VAYOATVCA	1592	9	12	86
VAYYFGLDV	1420	9	14	100
VSTLFOAVM	2632	9	12	86
VIOTVDFSL	1463	9	12	86
WAKHAWNFI	1766	9	12	86
YAOGYKVIL	1249	9	11	79
YAPTLWARM	2468	9	14	100
YSPGEINAV	2930	9	11	79
YSPGQNEF	2846	9	11	78
YSTYCKFLV	1298	9	12	86
YTNWDDOLV	1108	9	78	11
AAOGYKVVLV	1250	10	11	79
AAATLGEGAYM	1264	10	26	186
ASLRVTEAM	2787	10	12	86
ASSSASQSLA	2204	10	14	100
ATGNLPGCSF	165	10	13	93
CSFSFLAL	172	10	14	100
CTCGSSDYL	1128	10	11	79
DARVCACLWM	733	10	18	129
DSVIOQNTCV	1454	10	12	86
DTLTCGFADL	124	10	12	86
EAANLWRCBM	2237	10	24	171
ETAGARLVLV	1342	10	12	86
FAOULGMPFL	130	10	11	79
FTEAMTRSYA	2792	10	14	100
GAAPRALAHGV	146	10	11	79
GADTIAAGDGI	992	10	12	86
GAGVAGALVA	1661	10	86	11
GALVNGVCA	1695	10	79	11

HCY R58 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
GARLVLATA	1345	10	1.1	79
GAVONNNRLI	1916	10	1.4	100
GSGKSTKPA	1238	10	1.2	86
GTVLDOAFTA	1335	10	1.4	100
HSKKKQDELA	1400	10	1.0	100
IAFASRGHVV	1925	15	1.4	100
ISGQYLAGL	1774	10	1.4	100
ITRVESENKV	2250	10	1.2	86
ITSGSSNVSV	2818	10	1.4	100
ITYSTYGRFL	1296	10	1.1	79
KSTKVPAYA	1241	10	1.2	86
LADGSCSGGA	1305	10	1.1	79
LAEOPFKKAL	1729	10	1.2	86
LALPPRAYAM	806	10	1.2	86
LSPGALVIVV	1892	10	1.3	93
LSPRGSEPSW	88	10	1.1	79
LSHARPPWFM	3017	10	1.4	100
LSTLPGNPAI	1783	10	1.1	70
LTHPITKYIM	1842	10	1.6	114
NTGOTOTDF	1460	10	1.2	86
PAILSPGALV	1889	10	1.2	86
PALSTGJHL	888	10	1.2	86
PARUVFPL	2609	10	1.1	79
PSWQDAMKQ	1607	10	1.1	79
PTGSGKSTKV	1236	10	1.3	93
PTHYMPESDA	1936	10	1.2	86
PTLHGPTPL	1621	10	1.1	79
PTLWARMILM	2870	10	2.2	157
PTPLYFLGA	1626	10	1.3	93
DAETAGARLY	1340	10	1.2	86
QAPPSSWDM	1603	10	2.4	171
DATYCARAQ	1695	10	1.1	79
RAAKLQDCTM	2757	10	1.6	114
RAAVCTRGVA	1186	10	1.1	79
RAJAHGVRW	148	10	1.4	100
SASQSLAPS	2207	10	1.3	93
STKVPAYAA	1242	10	1.1	79
STWMLVGGW	1663	10	1.2	86
TAGARLWLA	1343	10	1.2	86
TARHTPVNSW	2852	10	1.1	79
TSCSSNVSA	2817	10	1.4	100
TSMLTDPSH	2177	10	1.3	93
TSTWVLVGGV	1662	10	1.2	86
TTIMAKNEVF	2589	10	1.1	79
TTLPALSTGL	605	10	1.1	79

HCV B58 Super Multif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy [%]
VAATLGEGAY	1263	10	1.4	100
VTGGERPSGM	1507	10	1.6	114
VTRHADIVPV	1138	10	1.1	79
WAQPGYPAWL	76	10	1.2	86
WARMILKTHF	2873	10	1.2	86
WARPDYNPL	2297	10	1.1	79
YAOQGYMLV	1249	10	1.1	79
YSPGEBINRVA	2930	10	1.1	79
YSPGQRFV	2648	10	1.1	79
YARALAHGVIV	147	11	1.1	79
AASLRVFTEAM	2788	11	1.2	86
AAVCTRGVAKA	1187	11	1.1	79
ASHLPIEIQGM	1717	11	1.1	100
ASQISAPSLSKA	2208	11	1.1	79
CARACQAPPSSW	1599	11	1.1	79
CSFSIFLLL	172	11	1.4	100
CTCGSSOLYLV	1128	11	1.1	70
CTRGVAKAVADF	1190	11	1.1	79
DARYGACLWMM	733	11	1.6	114
DLTTCGFAOLM	124	11	2.4	171
ETAGARLWVIA	1342	11	1.2	86
FADLIMGTYPLV	130	11	1.1	79
FSLJHSYSPGEB	2925	11	1.1	79
FTGLTHIDAHF	1567	11	1.3	93
FTTLPALSTGL	884	11	1.1	79
GADTAACGDI	992	11	1.2	86
GAGYAGALVAF	1061	11	1.2	86
GALVVGIVCA	1895	11	1.1	78
GAVQWMTLIA	1916	11	1.4	100
GSGKSTRVPA	1238	11	1.2	86
HSKKKCODELLA	1400	11	1.4	100
HSYSPGEBRN	2928	11	1.1	79
HTPVNSWGN	2855	11	1.2	86
ITRVESENKV	2250	11	1.2	86
ITSCSSNSVVA	2816	11	1.4	100
ITSYSTYKFLA	1286	11	1.1	79
KSTKVPAYAA	1241	11	1.1	79
LADGGSGSGAY	1305	11	1.1	79
LASYGAGVAGA	1857	11	1.1	79
LSNSLURHNM	2479	11	1.1	100
LSPGALWQWV	1692	12	1.2	86
LTCGFAOLMGT	126	13	1.3	93
LTSMLDPSH	2176	13	1.3	93
NAVAYYRGGLDV	1418	14	1.1	79
NTNHRPPOQW				

HCY_B58_Super_Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
PALSPGALVV	1888	1	12	86
PSVAAATLGFGA	1261	1	14	100
PTDPRRPSRL	109	1	12	86
PTHVVPESDAA	1836	1	12	86
PTLHGSPTPLLY	1821	1	11	79
PTPILYRLGAV	1626	1	13	93
GAETAGAARLVL	1340	1	12	86
QAPPSPSNDQAW	1603	1	11	79
QTVQESLQDPTF	1465	1	12	86
PSCPFHGRPH	55	1	13	93
SADLEWVTSWV	1655	1	11	79
SSASQSLAPSLL	2206	1	13	93
SSDLMLVTRHAA	1132	1	12	86
STWNLVGGVLA	1663	1	12	86
TARHPPVNSWV	2852	1	11	79
TSLTGDRDNQV	1050	1	12	86
TTWNLVGGVLL	1662	1	12	86
TTLPALSTGGLI	685	1	11	79
VATLQFGAYM	1283	1	11	106
VAGALVAFKVM	1864	1	14	100
VAVEPVFSDM	974	1	12	86
VAYOATVCAARA	1592	1	11	79
VAYYRGLOSVV	1420	1	14	100
VTSTWVVGCV	1661	1	12	86
WAQPGYPMPLV	76	1	11	86
WASHMLMTHFF	2873	1	12	86
YAOQGKYKVLV	1249	1	11	79
YATGQNLPGCSF	164	1	12	86
YTNWDDCLVGSW	1106	1	11	78

HCV B61 SuperMolif

Table XIV

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
AILSPGAL	1680	8	13	93
ALAHGVRY	150	8	14	100
ALGILQTA	1737	8	12	86
APTLWARM	2889	8	11	79
ADAPPISW	1602	8	12	86
AGYVYKLV	1221	5	11	79
AVAYYRQL	1419	8	14	100
AVCTAGVA	1188	8	11	76
AVQWHAIRL	1917	8	14	100
CLWMMILLI	739	8	12	86
CMSADLEV	1853	8	11	79
CDH-LFHW	1556	8	12	86
CYTOVDF	1462	8	12	86
DILAGIGA	1855	8	12	86
DLOGS/FL	279	8	12	86
DLMGKIPL	132	0	11	79
DLYNLPA	1883	8	11	78
DQAEATAGA	1339	8	12	86
EIPFYGKA	1377	8	13	93
EOPKQAL	1731	8	12	86
EVNTSTWV	1659	8	12	86
FISGICITL	1773	8	14	100
FPDLGVRY	2615	6	11	79
FRGGACSV	24	8	14	100
FQVHLHA	1226	8	12	86
GROYLAGL	1776	8	14	100
GLADLAVA	968	8	11	79
GPTLGVR	41	8	13	93
GSGGIGYY	26	6	14	100
GVAGGLVA	1863	8	12	86
GVAKANDF	1193	8	11	79
GVLAALAA	1670	8	12	86
GVRCCEK	2619	8	14	100
QWVCAAIL	1900	6	11	79
HAGPQEGA	1910	8	11	79
HVSPTHV	1933	8	12	86
ILGGWVAA	1816	8	12	86
ILGIGTVL	1331	8	12	86
ILSPQALV	1891	8	13	93
IMAKHEVF	2591	8	12	86
IPFYGKAI	1378	13	13	93
IPLVGAPL	137	8	11	79
IVDQVLY	701	12	12	86
IVPPOLGV	2613	8	11	79
IVGGVYL	30	8	13	93

HCY D62 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
KHALYDV	2625	8	12	88
KPARIIVF	2608	8	12	88
KOKAGLL	1734	8	12	86
KVPAAYAA	1244	8	9	79
LEANLLW	2235	8	11	12
LNTNGSW	414	8	12	86
LLAISCL	178	8	11	79
LLAPITAY	1030	8	12	86
LLADARV	729	8	12	100
LLYRIGAV	1629	8	13	93
LNGYFLV	133	8	11	78
LPALSTGL	687	8	14	100
LPQCSFSI	168	8	13	93
LPFRSPFL	37	8	13	93
LPVCDQHL	1553	8	13	93
LPYIEGM	1720	8	12	86
LODCTMVL	2761	8	12	86
LVAYCATV	1691	8	12	86
LYDILAGY	1853	8	11	79
LVGGVLA	1667	8	12	86
LVLNPSVA	1257	8	14	100
LVNLPLA	1884	8	11	79
LVTRHAOV	1137	8	12	86
LVVGIVCA	1897	8	11	79
LYVICESA	2773	8	11	79
MILMTHIFF	2878	8	12	86
MLTDPSHI	2179	8	14	100
MLLGWNA	1815	8	12	86
NIVDQYL	700	8	12	86
NLLYPOEM	2238	8	12	86
NPSVAAVL	1260	8	14	100
PLGGAAARA	143	8	11	79
PLLYFLGA	1628	8	13	93
PPPSVNDM	1605	8	12	86
PPSWDDMW	1606	8	11	79
PWHGCPL	4318	8	11	79
QVQGYL	29	8	13	93
QLRIPQA	336	8	12	86
QPEDLB	2808	8	11	79
QPGYWPPL	78	8	12	86
RLHCSAF	2918	8	12	86
RLNFPOL	2611	8	11	79
RLAPITA	1029	8	12	86
RLVVLATA	1347	8	12	86
RMAWDMMA	317	8	12	86

HCV_B62_Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
AMILLATHF	2875	8	12	86
APDQNPPL	2299	8	11	78
PGMGGN	2243	8	12	86
AVCEKHAL	2621	8	14	100
RVESENKV	2252	8	12	86
PGDHHY	210U	8	11	79
SIFLALL	175	8	14	100
SLDPTFTI	1470	8	14	100
SPQENRV	2931	8	11	79
SPQDRVF	2649	8	11	78
SQSLAPSL	2209	8	13	93
SVAAITLGIF	1262	8	14	100
TIMARKEV	2590	8	11	79
TIGFGAYM	1266	8	13	93
TLHGPTTL	1622	8	11	79
TLPGNPAAI	1785	8	11	79
TILWARMIL	2871	8	11	79
TPCSSESM	1876	8	12	86
TPGCPCPV	223	8	12	86
TQVQDESL	1464	8	12	86
TYCARADA	1597	8	11	79
VIDCTIV	1456	8	12	86
VLAALAY	1871	8	12	86
VLCFCYDA	1521	8	13	93
VLDQAEITA	1337	8	14	100
VLDEGQNY	157	8	12	86
VLNPSPVAA	1258	8	14	100
VLVQGMA	1668	8	12	86
VLVLPSPV	1256	8	14	100
VNGSSYGF	2639	8	11	79
VPESDAAA	1940	8	12	86
VQMMNRU	1918	8	14	100
VVATDAM	1439	8	11	79
VGIVVCAA	1898	8	11	79
VVTSTVVL	1860	8	12	86
WANRLLAF	1920	8	14	100
WPLLLL	798	8	12	86
WALVGM	1665	8	12	86
YLACLSL	1779	8	14	100
YPRWHY	616	8	14	100
YVPESDAA	1939	8	12	86
AILSGALV	1890	8	12	86
ALAHGIVRL	150	8	14	100
ALSTGLHL	688	8	12	86
ALVVGVVCA	1898	8	11	78

HCY_B62_SuperMolif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
APPSWIDOM	9	9	12	86
APTLWARM	9	9	11	79
AKGTYKAVL	9	9	11	79
ACPGYPPWPL	77	9	12	86
AVQNMNRU	19117	9	14	100
CMASADLEW	1653	9	11	79
DLCGSVFLV	279	9	11	79
DLEVNTSTW	1657	9	12	86
DLMGVPFLV	132	9	11	79
DVNLLPAI	1883	9	11	79
DLYVCEESA	2772	9	11	79
DLYVTRHAA	1134	9	12	86
DPODSQSSW	2410	9	11	79
DPRSRANL	111	9	12	86
EPPYGCAL	1377	9	13	93
EMGGNTTRV	2245	9	12	86
EVNTSTWVL	1658	9	12	86
FISQIYLA	1773	9	14	100
FLLAUSCL	177	9	12	86
FLLADARV	728	9	13	93
FQTSRGFRV	2646	9	11	79
GIGTVDQAA	1333	9	14	100
GIVYDQH	1552	9	13	93
GLADLAVAV	959	9	11	79
GLTHIDAHF	1569	9	13	83
GRGEGAVAV	1912	9	12	86
GPTPLVFL	1625	9	14	100
GQNGQVNL	28	9	13	93
GVAGALVAF	1863	9	12	86
GVLAALAY	1670	9	12	86
GVNYATGNL	161	9	11	79
GVVCEKMA	2618	9	14	100
GYMVEDGV	154	9	13	93
HJHQHIVDV	636	9	12	86
HJPIYEGDM	1718	9	11	79
HMYNFSQI	1769	9	13	93
HQHQDQY	698	9	11	79
HNGRGEAV	1910	9	11	78
ILAGYIAGV	1856	9	11	79
ILSPGALVV	1881	9	13	93
KVVLVLPSV	1255	9	14	100
LTSCSSNV	2815	9	14	100
LVFPDLGV	2612	9	11	79
LLFLLADA	728	9	14	100
LFNLGGW	1812	9	12	86

HCV B62 Super Multifano binding data

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
LLPRGRPL	26	8	13	83
LPAILSPGA	1888	9	13	93
LPALSTGLU	687	9	12	86
LCCEPFDV	2165	9	8	86
LPGCSFSIF	169	9	13	93
LYGGMLAAL	1884	2	12	86
LVLNPNSVAA	1257	14	14	100
LVLNLPLAIL	1884	9	11	78
LVTIHDADV	1137	9	11	79
LYGVIVCAA	1697	3	11	79
NILGGIVVAA	1815	3	12	86
NIRTGIVRAN	1282	9	11	79
NIVDQDLY	700	3	12	86
NLGRVDTL	118	9	12	86
NLPGCSFSI	168	9	13	93
NYDQDQVGV	1108	9	11	70
PLGGAARAL	143	9	11	79
PLLYFLGAV	1629	13	93	93
PPPSMDDMW	1665	9	11	79
PPVWIGCPL	2317	9	11	79
POPEYDLE	2807	9	11	78
PYQDQHLEF	1554	12	12	86
PYNSWLGNI	2857	9	14	100
QAVGIVYLL	29	13	13	93
QLSAPSLSKA	2210	9	11	79
QPEDDLQI	2808	9	11	79
OPGYMPWLY	78	9	12	86
OPRQPRQI	57	13	13	93
HLILAPITAY	1029	9	12	86
FLMLATHFF	2815	9	12	86
RVCEKHALY	2621	9	14	100
RYESENKVV	2232	9	12	86
RYAEGDGVY	156	9	12	86
SMLTDPISH	2178	9	14	100
SPGALWGV	1693	9	13	93
SPQEENRAV	2931	9	11	79
SPQOERVL	2649	9	11	79
SPRGGSAPSW	99	9	11	79
SVIDCNCIV	1455	9	12	86
TIMAHNEVF	2590	9	11	79
TJHGPTPL	1832	9	11	79
TLPAALSTGL	686	12	12	86
TLTCGGFADL	125	2871	78	93
TLWARMILM				
TPLYRLGA	1627			

HCV D62 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
TVLDQAEAT	1336	3	1.4	100
VIOTLTCGF	122	9	1.2	86
VLEDGVNVA	157	9	1.2	86
VLVDLILAGY	1832	9	1.1	78
VLVGGVGLAA	1658	9	1.2	86
VLVLNPSVA	1254	9	1.4	100
VQWMNLIA	1918	9	1.4	100
VQVVCAA	1698	9	1.4	100
WTSTWVWV	1860	9	1.1	78
WNRHFLAFA	1920	9	1.2	86
WNLVGGVLA	1665	9	1.2	86
YPLVGAPL	136	9	1.1	79
YLVATQATV	1590	9	1.2	86
YLVTFHADY	1136	9	1.2	86
YQATCARA	1594	9	1.3	93
YGDICGSV	276	9	1.2	86
YGVGEFL	637	9	1.3	93
YVPESSAAA	1938	9	1.2	86
AISPGALVV	1690	10	1.2	86
ALVVGIVCAA	1896	10	1.1	79
APPSPWQMW	1604	10	1.1	79
APTLWARMIL	2869	10	1.1	79
APCPYPMWLY	77	10	1.2	86
AVAYTGLGV	1419	10	1.4	100
AVCTRGVAKA	1168	10	1.1	79
AVGWNMRUA	1617	10	1.4	100
CLPKLGIVPL	2941	10	1.2	86
CVTQTVDFSL	1462	10	1.2	86
DILAGTIGAGV	1855	10	1.1	79
DLEVSTWV	1657	10	1.2	86
DLEVTVCDW	2817	10	1.3	93
DLSDGWSVTV	2412	10	1.1	79
DVNLLPAIL	1683	10	1.1	79
DOAETAGARL	1339	10	1.2	86
DMFPGGGCQ	21	10	1.2	86
ELITSCSSNV	2814	10	1.4	100
EQPKKALGL	1731	10	1.2	86
EVTSWVWLV	1658	10	1.2	86
GLSFSHSHY	2921	10	1.1	79
GLSTLPGNPA	1782	10	1.4	100
GLTHIDAHFL	1569	10	1.3	93
GPGEGLWQMM	1912	10	1.2	86
GVGCVGVL	28	10	1.3	93
GVCVTVYHGA	1091	10	1.1	79
GVRVCEKML	2619	10	1.4	100

HCV B67 Super Motif

Sequence	Position	Peptide No.	No. of Amino Acids	Sequence Frequency	Conservancy (%)
HONIVDION	808	10	11	11	79
LAGYGRGAV	1856	10	11	11	79
ILGGWVAQI	1858	10	12	12	86
IMAKNEVFCV	2591	10	11	11	78
IKYLAGLSTL	1777	10	14	14	100
INFDLGVRV	2613	10	11	11	79
KPTLHGFTPL	1620	10	11	11	79
KVDTLTCGF	121	10	12	12	86
KVLYLNPSVA	1255	10	14	14	100
LFNLLGHW	1812	10	12	12	86
LLPAISPGQ	1887	10	13	13	93
LMGYIPLVGA	133	10	11	11	79
LPAILSPEAL	1888	10	13	13	93
LPFGCSFSFL	169	10	13	13	93
LPREFSPRLGV	37	10	13	13	93
LPVQDQHLEF	1553	10	12	12	86
LVAYOATVCA	1591	10	12	12	86
LVQLAGYGA	1853	10	11	11	79
LVGGVLAALA	1667	10	12	12	86
LVGGVCAAI	1897	10	11	11	79
MLTDPSHITA	2179	10	14	14	100
NLPGCSFSIF	168	10	13	13	93
NPSVATLGF	1260	10	14	14	100
PITYSTGKF	1255	10	11	11	79
PLGGARALA	143	10	12	12	86
POPEYDLEI	2807	10	11	11	79
PVQDQHLEFW	1554	10	12	12	86
PNSWMLGNI	2857	10	14	14	100
PVYCFTRSPV	508	10	13	13	93
QUPCPPEPFDV	2164	10	12	12	86
OPEKGSGRKPA	2801	10	11	11	79
RLHGLSAPSL	2918	10	11	11	79
RLVFPDPOLGV	2611	10	11	11	79
RMAWDMAMNW	317	10	12	12	86
RALEDGVAYA	158	10	12	12	86
SUHYSYSPGEI	2956	10	11	11	79
SLTGRDIAKV	1051	10	12	12	86
SPGALVGIVV	1893	10	11	11	79
SOLSAFSLKA	2209	10	11	11	79
SOPRGFRPQ	56	10	13	13	93
SVAATLGFGA	1262	10	14	14	100
THGPTPFLY	1622	10	10	10	79
TLFNIIGGW	1811	10	12	12	86
TLPALSTGLI	686	10	11	11	79
YLTCGFADLM	125	10	12	12	88

HCY-B62 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
TPCTCQSSDL	1126	10	11	79
TPLYVRLGAV	1627	10	13	93
TPVNSVGNL	2856	10	12	86
TVDFLSDPFTF	1466	10	12	86
VIDLTCGFA	1122	10	12	86
VLAALAYCL	1871	10	12	86
VDOAETAGA	1337	10	12	86
VLNPSSVATL	1258	10	14	100
VLTTSGNLT	2737	10	11	79
VLVGGVLAAL	1666	10	12	86
VLVLNPSVAA	1256	10	14	100
VMSSSIQFQY	2638	10	11	79
VPESDAAARV	1940	10	12	86
VHMNRIIAF	1818	10	14	100
VIGWVCAIL	1898	10	11	79
WVLVGVLAAL	1665	10	12	86
YLKGSSGGPL	1165	10	12	86
YLPPRGRPPL	35	10	13	93
YLVTRHADVI	1136	10	11	79
YWGDLGSVF	218	10	12	86
ALVGVVYCAV	1898	11	11	79
APTGSSGKTKV	1235	11	13	93
APTLWARMILM	2869	11	11	79
AQAPPKPKDDM	1602	11	12	86
AVCTRGVAKAV	1188	11	11	79
AVOWNPNLIAF	1917	11	14	100
DILAGYGAQVA	1855	11	11	79
DLEWNTSTWNL	1657	11	12	86
DLGIVRCERMA	2617	11	13	93
DLMGVPVLGA	132	11	11	79
DLYLVTRHADY	1134	11	12	86
DOAETAGARLV	1339	11	12	86
DKKFFGGCON	21	11	12	86
EPKPKQKALGL	1731	11	11	100
FISGIGYLAGL	1773	11	14	86
FLADGCGSGA	1104	11	11	79
FPGGGCNGCN	24	11	14	100
FQYSPCRVEF	2646	11	11	79
GIOYLAGLSTL	1778	11	14	100
GLPVOOCHLEF	1552	11	12	86
GLSTLPGNPAI	1782	11	11	79
GPITPLYRLGA	1625	11	13	93
GPVYCFTPSPV	507	11	13	93
GVLAALAYCL	1670	11	12	86
GVRCVERMAY	2618	11	14	100

HCV JIG2 Super Multif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
GIVRLGGVNY	154	12	86	
H-LQHNDVQY	696	11	78	
HANMFSQIY	1789	11	83	
HONIVDQYDLY	898	11	79	
HYPGEAVQV	1910	11	79	
ILGGWVAQLA	1613	11	86	
ILGIGTVLDQI	1331	11	85	
ILSPGALWGV	1891	11	93	
KPARIYFPOL	2608	11	79	
KPTLHGPPL	1620	11	86	
KOKALQIOTA	1734	11	86	
KVDTLTCGFA	121	12	100	
KVLYVNSVAA	1255	14	100	
LIAFASRGNHV	1924	14	100	
LITSCSNSVSV	2815	14	100	
LIVFDLGVV	2612	11	79	
LLFLLLADARY	726	13	93	
LLFRILGGIWA	1812	12	86	
LLPAISPOAL	1887	13	83	
LLPRGRPHGLV	36	13	93	
LLSPRGSRPSW	97	11	79	
LLWPCERMGKQ	2240	12	86	
LPAILSPGALV	1888	12	86	
LPALSTGLHL	687	13	93	
LPGCSFSFLL	168	12	86	
LPGCGDLFEPV	1553	12	86	
LYGGVLAALAA	1667	12	100	
LVLMPSVATL	1257	14	100	
LVTRHADIVPV	1137	11	79	
LVGVIVCAIL	1897	11	78	
NIQGWWAQL	1815	12	86	
NITRVEEENKV	2249	13	93	
NLPAILSPGA	1688	11	11	
NPGCSFSFL	168	13	93	
PITYSTGKFL	1295	11	79	
PLEGEFGPOL	2403	13	93	
PMQFSYDTRCF	2687	11	79	
PPSHIDOMAKCL	1606	11	78	
PNSWIGNIIM	2857	12	86	
PVYCFIPSPV	508	13	93	
RKAVGGVYHRL	635	13	93	
RCEDGGMTRV	2243	12	88	
RCERKALYDV	2821	12	86	
SIFLLALLSCL	175	12	86	
SMLTDPSHTA	2178	14	100	

HCV R62 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
SPTHYPESDA	1835	1	12	86
SOURCEPENV	2163	1	12	86
SVAATLGGAY	1262	1	14	100
TGEGAYMSKA	1266	1	12	86
TLLFRILGGWV	1811	1	12	86
TPCTGSSSDLY	1126	1	11	79
TFGLPVCDHIL	1550	1	13	83
TPVNSHLGNII	2856	1	12	86
TVLQDAETAGA	1336	1	12	86
VLCECYDAGCA	1521	1	11	79
VLVDILAGYGA	1852	1	11	79
VLVGGVLAALA	1666	1	12	86
VPDEKGGRKPA	2600	1	11	79
VQWANRILFEA	1918	1	14	100
WVCAALRRHV	1901	1	11	79
WVLVGGVLAAL	1665	1	12	86
YLGSSSGPLL	1165	1	12	86
YLVAYOATVCA	1590	1	12	86
YQATVCARQD	1594	1	11	79
YVGOLCGSVP	276	1	12	86
YVPESDAAARV	1939	1	12	86
	428			

Table XV
IICV Δ Q1 Motif with Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A ⁰ 101
ASFQCSPY	166	26.0026	0	20	100
DNQVLSRKY	737	20.0255	10	10	80
FAAPFTQGY	631	20.0254	10	19	95
GFAAPFTQGY	630		11	19	95
GRETVLEY	140		9	15	75
GYSLNFMGY	579	2.0058	9	17	85
HTLWKGAGLY	149	1066.04	10	20	100
KOAFTEPPTY	653	20.0256	10	19	95
LLDTASALY	30	1066.01	9	17	85
LSDLVSAFVY	415	1090.07	10	19	95
LTFGRETVLEY	137		11	15	75
MHWYWMGPSLY	360	1030.01	10	17	85
MSTTDLEAY	103	2.0126	9	15	75
NSVVLRSKY	738	2.0123	9	10	90
PLDKGKPY	124	1147.12	9	20	100
PLDKGKPYY	124	1065.03	10	20	100
PTTGRTSLY	797	1090.09	9	17	85
SASFQCSPY	165		9	20	100
SLOVSAFVY	416	1000.02	9	19	95
STTDEAY	104		0	15	75
TGGRSLY	798	26.0030	0	17	85
WLSLUVSAFVY	414	26.0051	11	19	95
WMMWYWMGPS	359	1031.06	11	17	85
YPAALMPLY	640	19.0014	0	19	95
YSUNFMGY	580	26.0032	0	17	85
	2.5				

Table XVI UICVΔ03 Motif with Blinding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	Δ'0301
AACNWTRIGER	647	10	12	06	0.0003
AARALAHGVRA	147	10	11	79	
AAATLGFGA	1264	0	14	100	
AAATLGFGAY	1264	9	14	100	
AAVCTRGVA	1187	9	11	79	
AAVCTRGVAK	1187	10	11	79	
AAVCTRGVAKA	1107	11	11	79	
ACNWTRIGER	648	9	12	88	
ADGGCGSGA	1306	9	11	79	
ADGGCGSGAY	1306	10	11	79	
ADIVPVR	1142	0	12	06	
ADIVPVRAR	1142	9	11	79	
AFASERGNH	1926	0	14	100	
AGALVAFK	1065	0	12	06	
AGARLVLVA	1314	9	12	08	
AGARLVLVATA	1344	11	11	79	
AGLSTLPGNPA	1701	11	14	100	
AGVAGALVA	1062	9	12	06	
AGVAGALVAF	1062	10	12	86	
AGVAGALVAFK	1062	11	12	86	
AGWLLSPR	94	0	12	86	
AGWLLSPRSR	04	11	12	86	
AGYAGAGVA	1056	0	12	06	
AGYAGAGVAGA	1050	10	12	00	
ALGLLOTA	1737	0	12	06	
ALSTGLIH	609	0	12	06	
ALSTGLIH:H	609	10	12	06	
ALVVGVCA	1000	0	11	79	
ALVVGVCAA	1896	10	11	79	
ASLMAFTA	1793	0	11	79	
ASQLSAPSLK	2208	10	11	79	
ASQSLAPSLKA	2208	11	11	79	
ASRGNFVSPH	1928	11	12	06	
ASSASQSLSA	2204	10	14	100	
ATCNLPGCSF	165	10	13	93	
ATLGFGAY	1205	6	14	100	
ATLGFGAYMSK	1265	11	12	06	
ATRKTSER	46	0	0	79	
ATVCARAQAA	1596	9	9	79	
AVCTRGVA	1108	8	8	79	
AVCTRGVAK	1108	9	9	79	
AVCTRGVAKA	1108	10	10	100	
AVQWMNIFIA	1917	11	14	100	
AVQWMNIFIAF	1917	11	14	100	
CAAILRRH	1903	6	13	93	

MCV A03 Motif with Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'0301
CAWYELTPA	1530	9	1	79	
CGFADLMLGY	120	9	13	83	
CGNTLTCY	2742	6	11	79	
CGSSDLVLR	1130	11	11	79	
CGYRRCRA	2727	6	14	100	
CLRKLCVPPLR	2941	11	12	86	
CSFSIFLLA	172	9	14	100	
CSNVNSVA	2819	6	14	100	
CSNNVSAH	2819	9	12	86	
CTCGSSDLY	1120	9	11	79	0.0001
CTRGVAKA	1190	0	11	79	
CTRGVAKAVDF	1190	11	11	79	
CTWMNSTGFTK	555	9	11	79	
CWOPBKGR	2509	9	11	79	0.0000
CWOPERGGRK	2509	10	11	79	0.0011
CVTATVDF	1462	8	12	86	
DAHFSQTK	1574	9	14	100	0.0003
DOLVNCESA	2771	10	11	79	
DFSLOPTF	1460	0	14	100	
DGGCSGGA	1307	0	11	79	
DGGSGGGAY	1307	9	11	79	
DILICDECH	1310	9	12	86	
DILAGYGA	1055	6	12	86	
DILAGYGAAGA	1055	11	11	79	
DLGVRVCK	2017	9	13	93	0.0003
DLGWRVCEMA	2617	11	13	93	
DLMGYIPRLGA	132	11	11	79	
DLNLLPA	1803	8	11	79	
DLVNCESA	2772	9	11	79	
DLYLYTRH	1134	9	12	86	
DLYVTRHIA	1134	9	12	86	0.0003
DTLTGFA	124	8	12	86	
DWIPVARR	1143	8	11	79	
EAMTRYSA	2794	8	14	100	
DLYLYTRH	1134	8	11	79	
ECYDAGCA	1524	10	11	79	
ECYDAGCAY	1524	9	11	79	0.0004
EDLVNLPA	1882	9	14	100	
EGAVONMMR	1915	9	13	93	
EIPFYGKA	1377	8	12	86	
EMGGGNITR	2245	8	12	86	
ETAGARLVLIA	1342	11	12	86	
ETTMRSQVF	1207	9	12	86	0.0008
EVFOVQPK	2506	9	12	86	
FWOPERGGRK	2508	10	11	79	

ICV A3 Motif with Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'0301
FCVOPBKGGRK	2590	11	11	79	
FGAYMSKAH	1269	8	12	86	
FGCTWMNSTGF	1269	9	12	86	
FGYGIKDKVR	553	11	11	79	
FGYGIYLA	2554	9	12	86	0.0008
FISGIOYLA	1773	9	14	100	
FLADGGCGGQA	1304	11	11	79	
FLLADAR	728	8	14	100	
FSYDTRCF	2670	8	11	79	
FTEAMTRY	2792	8	14	100	
FTEAMTRYSA	2792	10	14	100	
FTGLTHIDA	1567	9	13	93	
FTGLTHIDAH	1567	10	13	93	
FTGLTHIDAHF	1567	11	13	93	
GAARLAH	146	0	11	79	
GAARLAHGVRA	146	11	11	79	
GAGVAGALVA	1061	10	12	86	
GAGVAGALVAF	1861	11	12	86	
GARHGVLA	350	8	12	86	
GALVGVCA	1895	10	11	79	
GALVGVYCA	1095	11	11	79	
GARLVLA	1345	8	12	86	
GARLVVLA	1345	10	11	79	
GAVOWMNR	1916	0	14	100	
GAVOWMNLIA	1916	11	14	100	
GAYMSKAH	1270	0	12	86	
GCAYWELTPA	1529	10	11	79	
GCSFSFLA	171	10	14	100	
GCTWMNSTGF	554	10	11	79	
GDDLVICESA	2770	11	11	79	
GDLCSSVF	2770	8	12	86	
GFAQLMGY	129	8	13	93	
GFGAYMSK	1268	8	12	86	
GFGAYMSKAH	1268	9	12	86	
GROSPQR	2645	10	12	86	
GSYDTRCF	2669	9	11	79	
GGAAABALA	145	8	11	79	
GGAAHLAII	145	9	11	79	
GCGGGAY	1300	9	11	79	
GGGGGGVY	26	10	14	100	
GGHYVOMA	935	8	11	79	
GGQVGGVY	27	9	14	100	
GGRLUFOH	1392	9	14	100	
GGRLUFCNSK	1392	14	11	100	

UCY A03 Motif with Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A.0301
GGRKPARLNF	2005	11	11	79	
GGVLAALAA	1669	8	12	86	
GGVLAALAA	1669	9	12	86	
GGVLAALAY	1669	10	12	86	
GGVLLPR	32	8	13	93	
GGVLLPRA	32	9	13	93	0.0003
GGWVAACQAA	1818	9	12	86	
GIGTVLDOA	1333	9	14	100	
GIYLPNRA	3037	8	11	79	
GLPVCOOH	1552	8	13	93	
GLPVCOOHLEEF	1552	11	12	96	
GLPVSAARH	1004	0	11	79	
GLPDLAYA	968	0	11	79	
GLSAFSULH	2921	0	11	79	
GLSFSLHSY	2921	10	11	79	0.0100
GLSTLPGNPA	1782	10	14	100	
GLTHIDAH	1569	0	13	93	
GLTHIDAHF	1569	9	13	93	
GSGKSTKVPA	1238	10	12	86	
GSGKSTKVPA	1230	11	12	86	
GSSQLYLVTTR	1131	10	12	86	
GSSDLYLVTTH	1131	11	12	86	
GSSYGEFY	2641	0	11	79	
GTFPINAY	2063	0	11	79	
GTYLDONETA	1335	10	14	100	
GVAGALVAF	1063	0	12	86	
GVAGALVAF	1063	9	12	86	
GVAGALVAFK	1003	10	12	86	0.3000
GVAKAIDF	1193	0	11	79	
GVCMWYH	1081	8	11	79	
GVCMWYHGA	1001	10	11	79	
GVGYLLPNA	3035	10	11	79	0.0014
GVLAALAA	1670	8	12	86	0.0046
GVLAALAY	1670	9	12	86	
GVRAFRKTSER	45	11	11	79	
GVRCCEKMA	2619	9	14	100	
GVRCCEKMA	2619	11	14	100	
GVRLDEGVY	154	11	12	86	
GVVCAAILR	1900	9	11	79	
GVVCAAILRH	1900	10	11	79	
GVVLLPRA	33	8	13	93	
GVVLLPRA	33	11	13	93	
HAQVIPVR	1141	6	79		
HAQVIPVR	1141	9	79		

HCV A03 Motif with Building Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A.0301
HADVIPVRAA	1141	10	1	79	
HAPTSGK	1234	0	1.4	100	
HAPTSGKSK	1234	11	1.3	83	
HGLSAFSUH	2920	9	11	79	
HGLSAFSUHY	2920	11	11	79	
HGPTPLLY	1624	0	11	79	
HGPTPLYR	1624	9	11	79	
HIDAHFLSQT	1572	11	14	100	
HILHAPTSGK	1232	10	12	86	0.5900
HILHONVQWY	696	11	11	79	
HILFHISK	1305	0	14	100	
HILFHISK	1395	0	14	100	0.0260
HILFHISRK	1395	10	14	100	0.0260
HMMWIFSGQY	1769	11	13	93	
HSKKKCDELAA	1400	10	14	100	
HSKKKCDELAA	1400	11	14	100	
HSYSRGEIR	2926	10	11	79	
HTPGCVPCVR	222	10	11	79	0.0004
HAGPSEGKA	1910	0	11	79	
IAFASRGNH	1825	9	14	100	0.0003
IDAHFLSQT	1573	10	14	100	
IDLTCGSE	123	8	12	86	
IDLTCGFA	123	9	12	86	
IFCHSKKK	1397	8	14	100	
IGTVLOOA	1334	0	14	100	
IGTVDOAETA	1334	0	14	100	
IIICDECH	1317	8	12	86	
ILAGYAGGVA	1056	10	11	79	
ILGGWVA	1016	8	12	86	
ILGGWVAACLA	1816	11	12	86	
ILGIGTVLOOA	1331	11	12	86	
IMARNEVF	2591	8	12	86	
ISGIVMLA	1774	0	14	100	
ITRVESENK	2250	9	12	86	0.0150
ITSCSSNNVVA	2816	11	14	100	
ITWGADTA	889	8	12	86	
ITWGADTA	909	9	12	86	
ITYSTYGK	1296	0	12	96	
ITYSTYGKF	1296	9	12	86	
ITYSTYGKFLA	1296	11	11	79	
IVDIOYLY	701	0	12	86	
IVFDLGV	2613	9	11	79	0.0036
IVGGVYLPR	30	10	13	93	0.0006
IVGGVYLPR	30	11	13	93	
KALGLLQTA	1736	9	12	86	

HCVΔ03 Multif with Blinding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A-0301
KCDELAAK	1404	8	12	86	
KFGYAKDVR	2553	10	12	86	
KGGRHLIF	1391	8	11	79	
KGGRHLIFCH	1391	10	11	79	
KGGRKPAR	2604	8	11	79	
KLGVPPPLR	2844	8	12	86	
KSTKVPAAA	1241	8	12	86	
KSTKVPAAY	1241	9	12	86	0.0008
KSTKVPAAYAA	1241	10	12	86	
KTKRANTRA	10	0	12	86	
KTKRANTRAA	10	9	12	86	
KTSERSOPRA	51	9	13	93	0.1600
KTSERSOPRGR	51	11	12	86	
KVDTLTCGF	121	10	12	86	
KVDTLTCGFA	121	11	12	86	
KVLVLNPSVA	1255	10	14	100	
KVLVLNPSVA	1255	11	14	100	
KVPAAYAA	1244	8	11	79	
LAGGCSGGA	1305	10	11	79	
LAGGCSGGAY	1305	11	11	79	
LAEOPKIC	1729	0	12	86	
LAEQFQIA	1720	9	12	86	
LAGYGAGVA	1057	9	11	78	
LAGYGAGVAGA	1057	11	11	78	
LCECYDAGCA	1522	10	11	79	
LDQNETAGA	1330	9	12	86	
LDQNETAGAT	1330	10	12	86	
LFILLADA	727	8	14	100	
LFILLADAR	727	9	14	100	
LFNILGGWVA	1013	10	12	86	
LFNILGGWVA	1013	11	12	86	
LTFFSPRA	290	8	11	79	0.0810
LGFGAYMSK	1287	9	12	86	
LGFGAYMSKA	1287	10	12	86	
LGFGAYMSKAH	1267	11	12	86	
LGGAARALA	144	9	11	79	
LGGAARALAH	144	10	11	79	
LGGMVAAQLA	1017	10	12	86	
LGIGTVLQDA	1332	10	13	83	
LGVRATRKG	44	8	12	86	
LGVRCEK	2618	8	14	100	
LGVRCEKMA	2618	10	14	100	
LIAFASRGNH	1924	10	14	100	
LIEANLLWRA	2235	9	12	86	0.0008

MCCYΔ03 Motif with Blinding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'0301
LIFCHSKK	1390	0	1.4	100	0.5400
LIFCHSKKK	1396	9	1.4	100	
LINTNGSMH	414	9	1.1	79	
LIVFPDGVIA	2612	10	1.1	79	0.0003
LLAPITAY	1030	8	1.4	100	
LLFLLLADA	726	9	1.4	100	0.0016
LLFLLLADAR	726	10	1.4	100	
LLFLNIGGWA	1012	11	1.2	86	
LLPAILSPGA	1887	10	1.3	93	0.0003
LLPARGGPR	36	6	1.3	93	
LLSPRGSR	97	0	1.2	06	
LMGYIPLVGA	133	10	1.1	79	0.0002
LSAFSLHRY	2022	9	1.1	79	
LSAPSLKA	2211	0	1.1	79	
LSNSLRLH	2479	0	1.2	06	
LSNSLRLRH	2479	9	1.2	06	0.0003
LSTGULH	690	9	1.2	06	
LSTLPGNPA	1703	9	1.4	100	
LTCGFADLMLGY	126	11	1.2	06	
LTDPHSHTA	2100	9	1.4	100	
LTHIDAHF	1570	6	1.3	93	
LTMUTDPSH	2176	10	1.3	93	
LVAYOATVCA	1591	10	1.2	86	
LVAYOATVCAH	1501	11	1.1	79	
LVDILAGY	1053	0	1.1	79	
LVDILAGYGA	1053	10	1.1	79	
LVGGVILAA	1667	0	1.2	06	
LVGGVILAAA	1667	10	1.2	06	
LVLNPSVA	1257	11	1.2	86	
LVLNPSVAA	1257	8	1.4	100	
LVVGVNCA	1887	8	1.1	79	
LVVGVNCA	1097	9	1.1	79	
LVVGWGA	2773	8	1.1	79	
MGFSYDTR	2668	8	1.1	79	
MGFSYDTRCF	2660	10	1.1	79	
MGSYYGFQY	2040	9	1.1	79	
MGYIPLVGA	134	9	1.1	79	
MILMTHFF	2076	0	1.2	86	
MLTDPSHTA	2179	10	1.4	100	
MSTNPKPDR	1	9	1.1	79	
MSTNPKPDRK	1	10	1.1	79	
NCGYRCPA	2726	8	1.1	79	
NCGYRCPRA	2726	8	1.1	79	
NCSYPGH	305	6	1.1	79	

LICV_A03 Motif with Blunting Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'0301
NFISGIOY	1772	0	1.4	100	
NFISGIOYLA	1772	10	1.4	100	
NCVCMVY	1000	0	1.1	79	
NGVCWVTVYH	1080	9	1.1	79	
NGVCWVYHGA	1000	11	1.1	79	
NILGAWWA	1815	8	1.2	86	
NILGGWAA	1815	9	1.2	86	
NITRVESENK	2249	10	1.2	86	0.0010
NIVDQYLY	700	9	1.2	86	0.0005
NLLPAILSPGA	1006	11	1.3	93	
NLPGECSFIF	160	10	1.3	93	
NTCVTQTVDF	1460	10	1.2	0.6	0.0010
NTNRPPQDKK	14	10	1.1	78	
NTNRPPQDKKF	14	11	1.1	79	
NTPGLPVQODH	1549	11	1.3	93	
PAISPGAA	1069	8	1.3	93	
PAISPGAA	600	9	1.2	86	
PALSTGLIHL	680	11	1.2	86	
POSGSMRL	1976	8	1.1	79	
PCTGQSSDLY	1127	10	1.1	79	
POLGVRYCEK	2616	10	1.3	93	
PGALVNGVCA	1094	11	1.1	79	
PGCSFSIF	170	9	1.4	100	
PGCSFSIFLLA	170	11	1.4	100	
PGCVPCVR	224	0	1.2	86	
PGEGAVWMMIA	1913	11	1.3	93	
PGEINRVA	2832	0	1.1	79	
PGERTPSGMF	1509	9	1.2	0.6	
PGGGGGGGY	25	11	1.4	100	
PGJPVQODH	1551	9	1.3	93	
PGYPMWLY	79	8	1.4	100	
PITYSTYKG	1295	9	1.1	79	
PITYSTYKG	1285	10	1.1	79	
PLGGAARA	143	6	1.1	79	
PLGGAARALA	143	10	1.1	79	
PLGGAARALAH	143	11	1.1	79	
PLLYRILGA	1628	8	1.3	93	
PMGFSYOTR	2667	9	1.1	79	
PMGFSYOTRCF	2667	11	1.1	79	
PSPWVGITDR	514	6	1.3	93	
PSVAATLGF	1261	9	1.4	100	
PSVAATLGFA	1261	11	1.4	100	
PSWQDMWKK	1607	8	1.1	79	
PTDCFRKH	507	9	1.3	93	
PTOPRARSA	109	9	1.2	0.6	0.0008

LLCYΔ01 Motif/Cyto Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'0301
PTGSGKSTK	1236	9	1.0	9.3	0.0002
PTHYVPESSDA	1936	10	1.2	8.6	
PTHYVPESSDA	1836	11	1.2	8.6	
PTLHQGPTPLLY	1621	11	1.1	7.9	
PTPILYALGA	1626	10	1.3	9.3	
PVCOCKLEF	1554	9	1.2	8.6	
PVWVGTTOR	516	9	1.3	9.3	0.0008
QAETGAGA	1340	8	1.2	8.6	
QATVCARA	1595	8	1.3	9.3	
QATVCARAQA	1595	10	1.1	7.9	
QIVGGVYLPR	20	11	1.3	9.3	
QIUTFSPI	209	0	1.2	0.6	
QLFTESPRFA	209	9	1.1	7.9	0.7500
QLLNRPOA	336	8	1.2	0.6	
QLSAPSILK	2210	6	1.1	7.9	
QLSAPSILKA	2210	9	1.1	7.9	
QTVDFSLDPTF	1465	11	1.2	8.6	
RAAVYCTRGVA	1186	10	1.1	7.9	
RAAVVCTRGVAK	1186	11	1.1	7.9	
RALAHGVIA	149	8	1.4	10.0	
RATRKTSER	47	9	1.1	7.9	
RGM-NSPTH	1930	9	1.2	8.6	0.0003
RGN-NSPTHY	1930	10	1.2	0.6	0.0003
RGPFLGVIPR	40	0	1.3	0.3	
RGPFLGVIPRA	40	9	1.3	0.3	
RGPFLGVIPTR	40	11	1.1	7.9	0.0120
RGRGROPIK	59	9	1.3	9.3	
RGSLLSPRA	1154	0	1.2	0.6	
RIGVAKAVDF	1192	9	1.1	7.9	
RIGVYRATR	43	8	1.1	7.9	0.9400
RIGVYRATRK	43	9	1.1	7.9	
RIGHGSASF	2918	8	1.2	8.6	
RIGHGSASFH	2918	11	1.1	7.9	
RILAFASR	1923	8	1.4	10.0	
RILAFASRGNH	1923	11	1.4	10.0	
RILNFPOLGVR	2611	0	1.1	7.9	
RILAPITA	1029	9	1.2	8.6	2.7000
RILAPITA	1029	0	1.2	8.6	
RILVILATA	1347	8	1.2	0.6	
RAMILMTHF	2075	8	1.2	8.6	
RAMILMTFF	2875	9	1.2	8.6	
RWYVGGVSH	635	9	1.4	10.0	
RWYVGGVSH	635	10	1.4	10.0	0.7200
RSOPCR	55	8	1.3	9.3	
RSOPCR	55	9	1.4	10.0	0.1800
RVCEKMLAY	2621				

HCVΔ03 Motif with Binding Information

Sequence	Position	Sequence	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A.0301
RVLEQGVNY	156	1174.17	9	1.2	86	0.0120
SAFSLHRY	156		10	1.2	86	
SASQSLAPSQK	2923		0	1.1	79	
SCSSANVSAH	2207		11	11	79	
SCSSNVSAH	2018		9	14	100	
SDLYLVTRH	2818		10	12	86	
SDLYLVTRH	1133		0	12	86	
SDLYLVTRH	1133		9	12	96	
SDLYLVTRH	1133		10	12	86	
SFSFLLA	173		6	14	100	
SGKSTKVPAA	1230		9	12	86	
SGKSTKVPAA	1239		10	12	86	
SGKSTKVPAA	1239		11	12	86	
SMLTDPASH	2170		0	14	100	
SMLTDPASHA	2170		11	14	100	
SSASOLSA	2206		0	14	100	
SSDLYLVTRH	1132		9	12	86	0.0003
SSDLYLVTRH	1132		10	12	86	0.0003
SSDLYLVTRH	1132		11	12	86	
SSNNSVSAH	2830		0	12	86	
SSSASQSLA	2205		9	14	100	
STGHLHLH	691		8	12	86	
STKVPAAV	1242		0	12	86	
STKVPAAV	1242		0	12	86	
STLPGNPA	1704		10	11	79	
STNPKPQR	2		0	14	100	
STNPKPQRK	2		9	11	79	
STWVLVGGVLA	1663		11	11	79	
STYGFELA	1299		8	12	86	
SVAAATLGCF	1262		0	14	100	
SVAAATLGCF	1262		10	14	100	
TAGARLVLVIA	1262		11	14	100	
TCGFADLMAY	1343		10	12	86	
TCGSSDOLY	127		0	13	93	
TCVQTQTVDF	1129		0	11	79	
TCVQTQTVDF	1461		9	12	86	
TDPARRSP	110		0	12	86	
TOPSHTA	2101		0	14	100	
TGEIPFYGK	1375		9	11	79	
TGEIPFYGK	1375		10	11	79	
TGLTHIDAH	1560		6	13	93	
TGLTHIDAH	1560		9	13	93	
TGLTHIDAHF	1568		10	13	93	

ICV Δ 03 Motif with Blinding Information

Position	Sequence	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A ⁺ 0301
166	TGNLPCCSF	9	1.3	9.3	
1237	TGSGKSTKVPAA	6	1.3	9.3	
1237	TGSGKSTKVPAA	11	1.2	9.6	
2590	TIMAKNEVF	9	1.1	7.9	
1266	TLGFGAYMSK	10	1.2	8.6	0.0810
1266	TLGFGAYMSK	11	1.2	8.6	
1622	TLHGPPTLLY	10	1.1	7.9	0.0890
1622	TLHGPPTLLY	11	1.1	7.9	
806	TLPALSTGLIH	11	1.1	7.9	
2071	TLWARMILMTH	11	1.1	7.9	
2017	TSCSSNVSYA	10	1.4	10.0	
2017	TSCSSNVSYA	11	1.2	9.6	
5.2	TSEASDFR	8	1.3	9.3	
5.2	TSEASDFR	10	1.2	9.6	0.0003
5.2	TSEASDFR	11	1.2	9.6	
1050	TSLTRGDK	0	1.2	9.6	
2177	TSMLTDPSH	9	1.3	9.3	0.0003
2589	TTIMAKNEVF	10	1.1	7.9	
1208	TTMRSVPF	6	1.2	8.6	
1597	TCARAOA	8	1.1	7.9	
1460	TVDFSLQPTF	10	1.2	9.6	
1336	TVLDQAEATA	9	1.4	10.0	
1536	TVLDQAEATA	11	1.2	9.6	
1263	VAAITLGFAQ	9	1.4	10.0	
1263	VAAITLGFAQ	10	1.4	10.0	
1064	VAGALVAF	9	1.2	9.6	
1004	VAGALVAFK	9	1.2	9.6	0.2400
1592	VAYQAVVCA	0	1.2	9.6	
1592	VAYQAVVCA	10	1.1	7.9	0.0006
1592	VAYQAVVCA	11	1.1	7.9	
1902	VCAAIRR	6	1.1	7.9	
1902	VCAAIRR	9	1.1	7.9	
2622	VCEKMLAY	8	1.4	10.0	
505	VCGPVYCF	6	1.3	9.3	
1555	VCGDHLF	8	1.2	9.6	
1169	VCTRGVAK	8	1.1	7.9	
1109	VCTRGVAKA	9	1.1	7.9	
1002	VCWIVVHQA	9	1.1	7.9	
1467	VDFSLOPFT	9	1.4	10.0	
1054	VDLAGYGA	9	1.1	7.9	
614	VDYPYRLWH	9	1.3	9.3	
614	VDYPYRLWH	10	1.3	9.3	
2597	VFCVQPEK	6	1.2	8.6	
2597	VFCVQPEKGR	7	1.2	8.6	
2614	VFPDLGVR	0	1.1	7.9	

MHCV A03 Motif with Binding Information

Sequence	Position	Sequence	Frequency	No. of Amino Acids	Conservancy (%)	A ⁰ 0301
VFTGLTHIDA	1566		10	13	93	
VFTGLTHIDA	1566		11	13	93	
VGDLOGSF	277		9	12	86	
VGGVLAALA	1668		9	12	06	
VGGVLAALA	1668		10	12	86	
VGGVLAALA	1668		11	12	86	
VGGVLAALA	1668		9	13	93	0.0003
VGGVTLPR	31		10	13	93	
VGGVMLPRA	3036		9	11	79	0.0007
VGYLLPNA	1098		10	11	79	
VGVVCAAILR	1099		11	11	70	
VGVVCAAILR	1099		9	12	06	
VIDLITCGF	122		9	12	06	
VIDLTGCF	122		10	12	06	
VIAALANY	1671		0	12	06	
VLCCEYDA	1521		0	13	93	
VLCCEYDAGCA	1521		11	11	79	
VLDQAEATA	1337		0	14	100	
VLDQAEATA	1337		10	12	86	
VLDQNETAGAR	1337		11	12	06	
VLEDGVNN	157		0	12	06	
VLEDGVNN	157		9	12	88	
VLNPSSVA	1250		0	14	100	
VLPSMLTDPSH	2176		11	13	93	
VLVDILAGY	1052		9	11	79	
VLVDILAGYGA	1052		11	11	79	
VLVGGVLA	1660		0	12	06	
VLVGGVLA	1660		9	12	86	0.0003
VLVGGVLAALA	1660		0	14	100	
VLVUNPSEVA	1256		9	14	100	
VLVUNPSSVA	1256		10	14	100	
VMGSSYSGF	2639		0	11	79	
VMGSSYCFQY	2639		10	11	79	
VIRHADIVPVR	1130		11	11	79	
VVCAAILR	1901		8	11	79	
VVCAAILR	1901		9	11	78	
VVCAAILR	1901		10	11	79	
VVGVNCAA	1098		0	11	79	
VVGVNCAA	1098		11	11	79	
VVGGVCAAILR	517		0	13	93	
WAGWLLSPR	93		9	12	06	
WAKHMMWF	1766		8	12	86	
WAQPGYPWPLY	76		11	12	86	
WARMILMTH	2073		0	12	86	
WARMILMTH	2873		10	12	86	
WARMILMTHFF	2873		11	12	86	

HCVΔ03 Motif with Blinding Information

Position	Sequence	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'0301
107	WGPTOPRRA	0	12	86	
107	WGPTOPRRA	9	12	86	
107	WGPTOPRRA	11	12	86	
96	WLLSPRGSR	9	12	86	0.0008
1920	WMNRLLAF	8	14	100	
1920	WMNRLLAF	9	14	100	0.0003
1920	WMNRLLAFASR	11	14	100	
557	WMNSTGFTK	9	11	79	0.0530
1695	WLVLGGVLA	9	12	86	
1695	WLVLGGVLA	10	12	86	
164	YATGNLFGCSF	11	12	86	
1526	YDAGCAWY	0	11	70	
1315	YDNCDECH	10	12	86	
1060	YGAIVVAGA	0	12	86	
1060	YGAIVVAGA	11	12	86	
2644	YGRGYSPOOR	10	11	70	0.0054
35	YLLPRPGPR	9	13	83	
1590	YLVAYOATVCA	11	12	86	
2930	YSPGEINRVA	6	11	79	
2930	YSPGEINRVA	10	11	79	
2648	YSPGCRVEF	9	11	79	
1298	YSTYGKFLA	9	12	86	
276	YVGQLOGSVF	10	12	86	
637	YVGVB-R	0	14	100	
1930	YVPESDAA	0	12	86	
1930	YVPESDAA	9	12	86	0.0003
1939	YVPESDAA	10	12	86	
1939	YVPESDAA	3			

Table XVII
IICV All Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'11101
AACNWTRGER	647	10	1.2	86	0.0140
AARALAI-GVR	147	10	1.1	79	
ATLGF-GAY	1264	9	1.4	100	
AAVCTRGVAK	1107	10	1.1	79	
ACNWTRGER	640	9	1.2	86	
ADGGCGGGAY	1306	10	1.1	79	
ADVIPVYR	1142	0	1.2	86	
ADVIPVYR	1142	9	1.1	79	
AFASRGNH	1926	8	1.4	100	
AGALVAFK	1005	0	1.2	86	
AGVAGALVAFK	1062	11	1.2	86	
AGMLSPR	94	0	1.2	86	
AGWLSFRGSA	94	11	1.2	86	
ALSTGLIH	689	0	1.2	86	
ALSTGLIH	609	10	1.2	86	
ASQSLAPSILK	2200	10	1.1	79	
ASRGNHVSPTH	1928	11	1.2	86	
ATLGF-GAY	1265	0	1.4	100	
ATLGF-GAYMSK	1205	11	1.2	86	
ATKTSER	48	0	1.1	79	
AVCTRGVAK	1100	9	1.1	79	
CMLLRAH	1903	0	1.3	93	
CGFAQMAGY	128	9	1.1	79	
CGNTLTCY	2742	0	1.1	79	
CGSSPLYLVTR	1130	11	1.2	86	
CLRLGPPLR	2941	11	1.1	79	
CNSIYPGH	304	9	1.1	79	
CNWTRGER	049	0	1.2	86	
CSANVVAH	2018	9	1.2	86	
CTCGSSDLY	1128	9	1.1	79	
CTWMNTGFTK	555	11	1.1	79	0.0250
CYOPERGGR	2599	9	1.1	79	
CYOPERGGR	2599	10	1.1	79	
DAHFLSQTK	1574	0	1.4	100	
DGGCGGGAY	1307	9	1.1	78	
DNCDECH	1316	9	1.2	86	
DGVVRCER	2617	9	1.3	93	
DVVLTRH	1134	8	1.2	86	
DVPLVRR	1143	8	1.1	79	
ECYDAGCAY	1524	10	1.1	79	
EGAVQWMNR	1815	9	1.4	100	0.0014
EMGEGNTR	2245	8	1.2	86	
EVFCVQPEK	2596	9	1.2	86	0.0270
FCVOPKGGR	2598	10	1.1	79	
FOVOPKGGR	2598	11	1.1	79	

ICV ALM Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'1101
FGAYMSKAH	1269	9	12	66	
FGYGAKDVA	2554	9	12	66	0.00005
FLLADAR	728	8	14	100	
FTEAMTRY	2782	8	14	100	
FTGLTHIDAH	1567	10	13	93	
GAARALAH	146	6	11	79	
GAARALAHGVR	146	11	11	79	
GAVQMMAR	1916	9	14	100	
GAYMSKAH	1270	8	12	88	
GFAOLMGY	129	0	13	93	
GFGAYMSK	1260	0	12	06	
GEGAYMSKAH	1260	10	12	06	
GFGYSPQR	2645	9	11	79	
GGAARALAH	145	9	11	79	
GGCCSGAY	1308	0	11	79	
GGCGANGGVY	26	10	14	100	
GGGQSGCVY	27	9	14	100	
GGHLFLICH	1382	9	14	100	0.00001
GGHLFLIFHSK	1392	11	14	100	
GGVLAALAAAY	1669	10	12	86	
GGVLLPRA	32	9	13	93	0.00010
GGVYLPRR	32	9	13	93	
GYLPLPR	3037	0	11	79	
GLPVOQH	1652	0	13	93	
GLPVSAAR	1004	0	11	79	
GLSAFSLH	2921	0	11	78	
GLSAFSLHSY	2021	10	11	79	
GLTHIDAH	1569	0	13	93	
GNHVSPTH	1931	0	12	86	
GNHVSPTHY	1931	0	12	86	
GNTRVSEENK	2248	11	12	86	
GSSDLVLYTRH	1131	10	12	86	
GSSDLVLYTRH	1131	11	11	79	
GSSYGFQD	2641	0	11	79	
GTFFPINAY	2063	0	11	79	
GVAGALVAFK	1863	10	12	86	1.4000
GVCMWTH	1081	8	11	79	
GVGYLPLR	3035	0	11	79	0.0140
GVLAALAAAY	1670	9	12	86	0.0110
GVRAFKTSE	45	11	11	79	
GVRCVCEKALY	2619	11	14	100	
GVRLDEGQY	154	12	12	86	
GVVCAALRLR	1900	9	11	79	
GVVCAILRLR	1800	10	11	79	
GVVCAILRLR	1900	11	11	79	

MCV All Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'1101
GYVLPRA	33	0	13	93	
GYLPARGRA	33	1	13	83	
HADIVIVRA	1141	9	11	79	
HADIVVYRA	1141	9	11	78	
HADIVVRRR	1141	10	11	79	
HAPTSGSK	1234	8	14	100	
HAPTSGSKSTK	1234	11	13	93	
HGLSFSLH	2920	9	11	79	
HGLSFSLHSY	2920	11	11	79	
HGPPIPLY	1624	0	11	79	
HGPPIPLYR	1624	9	11	79	
HIDAHFSOTK	1572	11	14	100	
HJHAPGSGK	1232	10	12	86	0.0024
HJHONIVDQY	696	11	11	79	
HJFCSK	1305	0	14	100	
HJFCSKK	1305	9	14	100	0.0000
HJFCSKKK	1395	10	14	100	0.0002
HMWNFGIQY	1769	11	13	93	
HYSPEINR	2920	10	11	79	
HTPGCVCVR	222	10	11	79	
IAFASGNH	1925	9	14	100	0.0012
IDAIFLSOTK	1573	10	14	100	0.0003
IFCHSKKK	1397	9	14	100	
IIICDECH	1317	8	12	86	
INTNGSWH	415	0	11	79	
ITVWESENK	2250	9	12	86	
ITYSTGK	1200	0	12	86	
IVDQKLY	701	0	12	86	
IVFPDLGVR	2613	9	11	79	0.0044
IVGGYLLPRA	30	10	13	93	0.0066
IVGGYMLPRA	30	11	13	93	
KCDELAAK	1404	8	12	86	
KFGVGAOKVR	2553	10	12	86	
KGGFHLFCH	1381	10	11	79	
KGGFKPAR	2804	8	11	79	
KLGVPPRL	2844	6	12	86	
KNEVFCOPEK	2594	11	11	79	
KSTKVPAY	1241	9	12	86	0.0001
KTAKANTNR	10	0	12	86	
KTAKANTNR	10	9	12	86	0.0100
KTSESRSPRA	51	9	13	93	0.0640
KTSESRSPRGR	51	11	12	86	
LAQGEGSGAY	1305	11	11	79	
LAEOFK	1729	9	12	86	
LDQAEATGAR	1338	10	12	86	

HCV 3' UTR Motif With Blinding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A 1101
LFLLLADAR	727	9	1.4	100	
LFTFSPPR	290	0	1.1	79	
LGFGAYMSK	1267	9	1.2	86	0.2800
LGFGAYMSKAH	1267	11	1.2	86	
LGGAARALAH	144	10	1.1	79	
LGVRATRK	44	8	1.2	86	
LGVRACEK	2618	0	1.4	100	
LIAFASGNH	1924	10	1.4	100	
LIEANLWWR	2235	9	1.2	88	0.0005
LIFCHSKKK	1596	0	1.4	100	
LIFCHSKKK	1300	9	1.4	100	0.1900
LINTNGSMWH	414	9	1.1	79	
LINPDLGVR	2612	10	1.1	79	0.0001
LLAPITAY	1030	U	1.4	100	
LLFLLLADAR	726	10	1.4	100	
LLPRGGPR	36	0	1.3	93	
LLSPRGSSA	97	8	1.2	86	
LSAFSLSHSY	2922	9	1.1	79	0.0002
LSNSLRLH	2479	0	1.2	86	
LSNSLRLH	2479	9	1.2	86	0.0001
LSTGLHLH	690	9	1.2	86	
LTGGFAADMGY	126	11	1.2	93	
LTSMLTPSH	2176	10	1.3	93	
LVAYQATVCAR	1591	1	1	79	
LVDILAGY	1053	0	1	79	
MGFSTYDIIH	2680	0	1	79	
MGSSYGFQY	2640	9	1	79	
MNRLIAFASR	1921	10	1.4	100	
MNSTGFTK	550.	0	1	79	
MSTNPKPOR	1	9	1	79	
MSTNPKPKR	1	10	1	79	
NGTYFPER	2726	9	1	79	
NCISYPGH	305	8	1	79	
NFISGQY	1772	0	1.4	100	
NGVCMVY	1080	8	1	79	
NGVCMVYH	1000	9	1	79	
NITRVESENK	2249	10	1.2	86	0.0062
NIVDQQLY	700	9	1.2	86	0.0140
NTNARPOVK	14	0	1	79	0.0007
NTPGLPVCOOH	1549	11	1.3	93	
PALSTGLHLH	608	9	1.2	86	
PCSGSMR	1976	8	1.1	79	
PCTGSSDLY	1127	10	1	79	
PDGLGVACEK	2616	0	1.3	93	

LICV All Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'1101
PGCVFCVA	224	6	12	0.6	
PGEAGAVMNN	1913	11	13	0.3	
PGGGGNGGVY	25	11	14	100	
PGLPIQODH	1551	9	13	93	
PGYPMFLY	79	0	14	100	
PITYSTYKG	1295	9	11	79	
PLGAAARALAH	143	11	11	79	
PMGFSYDTTR	2667	9	11	79	
PNIRTGVR	1281	8	13	93	
PSPVVVGTTDR	514	11	13	93	
PSWDDAMWK	1607	0	11	79	
PTDCPRK	507	0	13	93	
PTDPRRSR	109	9	12	0.6	0.0005
PTGSGKSTK	1238	9	13	93	0.0001
PTLHGPTPLLY	1621	11	11	79	
PVWNGTDR	516	9	13	93	0.0005
QAEFTAGAR	1340	8	12	86	
QIVGGVYLPR	28	11	13	93	
QUFFSPR	209	0	12	86	
QLFTFSPAR	289	9	11	78	0.0330
QLSAPSLK	2210	9	11	79	
QNIWQDQY	699	0	11	79	
QNIWQDQYLY	699	10	11	79	
RMAAVCTGIVAK	1100	11	11	79	
RALAHGVR	149	0	14	100	
RATRKTSER	47	9	12	0.6	0.0001
RCNNIVSPTH	1930	9	10	12	0.6
RGHNHSPTHY					
RGPRPLGVR	40	8	13	93	
RGPRLGIVATR	40	11	11	79	
RGRHOPK	59	9	13	93	0.0017
RGSLSPSR	1154	9	12	86	
RLGVPRTR	43	8	11	79	
RLGVPRTRK	43	9	11	79	0.0280
RHGCLSAFSLH	2918	11	11	79	
RUIAFASR	1923	9	14	100	
RUIAFASRGNH	1933	11	14	100	
RUVFDFLGVR	2611	11	11	79	
RLLAPATAY	1029	9	12	86	0.0270
FMVNGGVETH	635	9	14	100	
RMVSGVENVF	635	10	14	100	0.0200
RNTNHRPQDK	13	11	11	79	
RSOPPGRR	55	0	13	93	0.5000
RVCEKMLAY	2621	9	14	100	
RWLEDGVNY	150	12	12	86	0.0068

HCVΔUU Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A'1101
SAFSLHSY	2923	8	11	79	
SASQSLAPSILK	2207	11	11	79	
SCSNSAVSAH	2818	10	12	86	
SDLYLVTIAH	1133	9	12	86	
SDLYLVTIH	1133	9	12	86	
SGKSTKVPAAY	1239	11	12	86	
SMLTDPSH	2178	9	14	100	
SNSLPHH	2400	9	12	86	
SSDLYLVTIAH	1132	9	12	86	
SSDLYLVTIH	1132	10	12	86	
SSNSAVSAH	2020	0	12	0	
STGLILH	691	0	12	86	
STKVPAAY	1242	8	12	86	
STNPKPOR	2	0	11	79	
STNPKPORK	2	9	11	79	
STNPKPORKTK	2	11	11	79	
SVAATLGFAY	1262	11	14	100	
TCGFAQDQMGY	1277	10	13	83	
TCGSSDLY	1129	0	11	79	
TDPRSSR	110	8	12	86	
TGEIPFGIK	1375	9	11	79	
TGLTHIDAH	1568	9	13	93	
TGSGKSTIK	1237	0	13	93	
TLGFGAYWISK	1208	0	12	06	
TLHGPTPLLY	1622	0	11	79	
TLHGPTPLYR	1622	11	11	79	
TLPA1STGLIH	000	11	11	79	
TLWARMILMTH	2071	11	11	79	
TNPKPOR	3	8	11	79	
TNPKPORKTK	3	10	11	79	
TNPKPORKTKR	3	11	11	79	
TNRPQDVK	15	0	12	86	
TSCSENNSVVAH	2817	11	12	86	
TSERSOPR	52	0	13	93	
TSERSOPRGR	52	10	12	86	
TSERSOPRGR	52	11	12	86	
TSLTGADK	1050	0	12	86	
TSMLTDPSH	2177	9	13	93	
VAATLGFAY	1263	10	14	100	
VAGLVAFK	1864	8	12	06	
VAYQATVCAAR	1592	10	11	79	
VCAAILRR	1802	8	11	79	
VCAILRR	1902	9	11	79	
VCEKMLY	2622	6	14	100	
VCTRGVAK	1169	0	11	79	

MCY ALL Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A11101
VDYPYFLWH	614	9	13	93	
VDYPYFLWHY	614	10	13	93	
VFCIGPEK	2597	0	12	06	
VFCIGPEKGR	2597	11	11	79	
VFPDLGVR	2614	8	11	78	
VFTGLTHIDAH	1566	11	13	93	
VGGVLALAAV	1668	11	12	86	0.0019
VGGVTLPR	31	9	13	93	
VGGVYLPRA	31	10	13	93	
VGYLLPNN	3036	9	11	79	0.0100
VGWVCAAILR	1099	10	11	79	
VGWVCAILRR	1099	11	11	79	
VLAALAY	1671	8	12	06	
VLDQAEATGAR	1337	11	12	88	
VLEDGVNY	157	0	12	86	
VI.TSMLTDPSH	2175	11	13	93	
VLDVILAGY	1052	9	11	79	
VMGSSYGFQY	2638	10	11	79	
VTRHADVIPVR	1138	11	11	79	
VVCAAILR	1001	9	11	79	
VVCAAILRAH	1001	9	11	79	
VVCAAILRAH	1801	10	11	79	
VVGWVCAAILR	1000	11	11	79	
WVGTTON	517	0	13	93	
WAGWLSPH	93	0	12	06	
WADPGYPWPL	76	11	12	86	
WARMILMTH	2073	0	12	86	
WGPTOPRA	107	0	12	06	
WGPTOPRAA	107	9	12	86	
WGPTDPHPRSR	107	11	12	86	
WLLSPFGSR	96	9	12	86	0.0005
WMNRLIAFASR	1920	10	14	100	0.0010
WMNSTGFTK	557	9	11	79	
WNFESIDY	1771	9	14	100	
YDGCAYW	1526	0	11	79	
YDNCDECH	1315	10	12	86	
YGFQYSPQR	2644	10	11	79	
YLPRIGPRA	35	9	13	93	0.0005
YSPQEINR	2930	0	11	79	
YVGGVDR	637	0	14	100	
YVPESDAAR	1939	10	12	86	0.0001
	311	3			

Table XVIII
ICYΔ24 Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*2401
AWDMMIMNW	319	0	12	0.6	
AYAAQGYKVL	1248	10	11	79	0.0009
AYYRGIDVSV	1421	11	14	100	
CYDAGCAW	1525	0	11	79	
CYDAGCAWYEL	1525	11	11	79	
DFSLDPIFTI	1468	0	14	100	
FWAKHMMNF	1765	10	14	100	
FWAKHMMNF	1765	9	12	06	6.9000
GFADLWGYI	129	9	13	93	
GFADLWGYIPL	129	11	11	70	
GFSYDTRCF	2609	9	11	79	
GWTLALPI	1027	8	11	79	
GYGAGVAGAL	1059	10	12	06	0.0003
GYIPLVGAPL	135	10	11	70	0.0057
GYARRGRASGVN	2720	11	12	06	
HMWNFISGI	1769	9	13	93	
IFLLALSLCL	176	10	12	06	
IMAKHEVF	2591	8	12	86	
KPGGCCQ	23	8	13	93	
LFNLLGGW	1813	8	12	86	
LWARMILNTHF	2872	11	12	86	
LWHDENSGAN	2241	10	12	86	
LYLVTAHAOWV	1135	11	11	79	
MWNFTSGI	1770	0	14	100	
MWNFTSGIYL	1770	11	14	100	
MVGGVIEHRL	636	10	13	93	0.0270
NFISGQYL	1772	9	14	100	0.0170
PMGFSYDTRCF	2667	11	11	79	
QFKOKALGL	1732	9	12	86	
QFKOKALGL	1732	10	12	86	
QWMNRLLAF	1919	9	14	100	
QYLAGLSTL	1778	9	14	100	0.0400
QYSPGQRYEF	2647	10	11	79	0.0180
QYSPGQRYEF	2647	11	11	79	
FIMAWDMMMW	3117	10	12	86	
RMILMTHF	2075	0	12	86	
RMILMTHF	2075	9	12	86	
AMYVGAEVFLR	635	11	13	93	
SFSIPLL	173	9	14	100	
SFSIPLL	173	10	14	100	0.0041
SMLTFSHI	2778	9	14	100	
SWDQAMMKL	1668	9	11	79	
SYLGESGGPL	1164	12	86		
TWNINSTGF	556	11	8		

HCV_A24 Motif With Blinding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A ²⁴⁰¹
TMVGGM	1664	9	1.2	0.6	
TSTYGF	1297	0	1.3	0.3	
TSTYGRL	1297	9	1.2	86..	0.0230
VFTGLTHI	1566	8	1.3	0.3	
VMGSSYGF	2639	8	1.1	79	
VLLPREGPRL	34	11	1.3	9.3	0.0016
WMNRLIAF	1920	8	1.4	100	
YRGLDYSM	1422	10	1.4	100	
	53	2			

Table XIX a
LUCY DIL-SUPER-MOLIR

Core Sequence	Core Freq.	Core Consistency (%)	Exemplary Sequence	Exemplary Frequency	Position In HCV Poly-protein
FGAYNSKAI	12	80	FLGAYNSKAI	1266	16
FGCTWMAST	12	86	CHNGCTWMAST	550	79
FIKQVAGLL	12	86	AGFQKQVAGLLQIA	1730	86
FILALLSCL	12	86	FSFLALLSCLVLP	114	43
FIDGLRWC	11	79	LYFIDGLRWC	2612	79
FOVAKHLAP	12	86	POFDQVAKHLAP	1225	43
FRVAKHTQ	12	86	VDIFRVAKHTQ	1182	43
FSEFLALL	14	100	GCSEFLALLSCL	121	50
FSDFPIFTI	14	100	TVDFSDPIFTI	1466	86
FTEANRVS	14	100	UNFEATENRVS	2769	73
FTFSPPAVG	13	93	WYCFSPPPAVG	569	13
FTTLPALST	11	79	PCSETTLPALSTG	681	93
FWNGKAMMF	12	86	LEWYKAMWAMMF	1702	64
IANFESQI	14	100	LTIDWYKIANFESQI	1570	21
IDCNICVTO	12	95	DSMDCNICVTOID	1454	50
IDLTGCFEA	12	98	GRMIDLTCGCFEA	120	86
IEANLWQH	12	86	AUIEANLWQH	2233	86
IFULALLSC	14	100	SFSIFULLSCLIV	173	50
ILGQWAAQ	12	86	LFPLGQWAAQ	1613	43
ILGQYVLO	12	95	STTILGQVLODNE	1326	57
ILPAPVQPO	11	79	CANLQVWPGQDIA	1903	57
ILSPQALWV	12	93	LPALSPQALWV	1086	79
INAYTTCG	11	86	TEPRAATTCG	2064	57
IRFLVQVPLQ	11	79	MDYFLVQVPLQAA	134	10
ITIVNESENK	12	86	CGCTTIVNESENK	2247	71
ISCSASSNS	14	100	LEUTISCSASSNSV	2013	71
ITPFLAALC	12	79	ALDIPFLAALC	2810	79
ITVFLVQVPLQ	12	86	QVQVFLVQVPLQ	1669	6
ITVIVVQVPLQ	11	79	QVFLVQVPLQ	1302	57
ITVIVVQVPLQ	11	100	ITVIVVQVPLQ	1777	14
ITVIVVQVPLQ	11	79	VVLLFLDARVCS	1654	10
ITVIVVQVPLQ	12	86	LVVLLFLDARVCS	1348	71
ITVIVVQVPLQ	12	66	DESLVITV	1466	64
ITVIVVQVPLQ	12	66	ITVIVVQVPLQ	1335	6
ITVIVVQVPLQ	13	93	QVFLVQVPLQ	2810	13
ITVIVVQVPLQ	12	86	SYDQEVITV	1655	93
ITVIVVQVPLQ	14	100	SYDQEVITV	11	6
ITVIVVQVPLQ	11	79	WVLLFLDARVCS	724	4
ITVIVVQVPLQ	12	86	FRIGGGWVWQVPLQ	1614	29
ITVIVVQVPLQ	12	86	TTQVQVPLQ	1329	8
ITVIVVQVPLQ	12	66	CPVQVQVPLQ	1329	57
ITVIVVQVPLQ	12	66	QVQVQVPLQ	41	64
ITVIVVQVPLQ	14	100	PRQVQVPLQ	2815	6
ITVIVVQVPLQ	11	79	PRQVQVPLQ	2916	11
ITVIVVQVPLQ	11	79	KTUQVQVPLQ	1620	13
ITVIVVQVPLQ	12	86	WVLLFLDARVCS	694	19
ITVIVVQVPLQ	11	79	WVLLFLDARVCS	2924	71
ITVIVVQVPLQ	14	100	WVLLFLDARVCS	1921	12
ITVIVVQVPLQ	12	66	WVLLFLDARVCS	2233	7
ITVIVVQVPLQ	14	100	GRVFLVQVPLQ	1393	14
ITVIVVQVPLQ	14	100	QVFLVQVPLQ	2812	100
ITVIVVQVPLQ	11	79	SPFLVQVPLQ	1716	13
ITVIVVQVPLQ	10	66	YVFLVQVPLQ	723	5
ITVIVVQVPLQ	13	93	YVFLVQVPLQ	1606	1
ITVIVVQVPLQ	12	86	YVFLVQVPLQ	725	29
ITVIVVQVPLQ	13	93	YVFLVQVPLQ	164	64
ITVIVVQVPLQ	13	93	YVFLVQVPLQ	1664	1

HCV Polyproteins

Core Sequence	Core Freq.	Core Consistency (%)	Exemplary Sequence	Position in HCV Polyprotein	Exemplary Sequence Frequency	Exemplary Sequence Consistency (%)
LAQYPLQD	11	79	FADQDPLVQAPR	130	11	79
LNPSSVAAIL	14	100	VLVNPSSVAAILQFA	1256	14	100
LPATLSPGIA	13	93	VALPALSQDPLV	1885	11	79
LPASLSTOLI	12	86	FTPLALSQDPLV	601	11	79
LPFTGPGFLH	13	93	VLWLPFTGPGFLH	34	13	79
LRDLAVAE	11	79	INGLIDQIAVAEVYV	966	4	29
LNKLNQPPH	12	86	ASCDLNQKQPLDHW	2939	7	50
LSAFSCHSYK	11	79	LFQLSAFSYKISYSPG	2919	11	79
LSAFSLSKAT	11	79	ASQ1SAPSLKATCTT	2206	7	50
LSNLSPH	12	86	IVASLNSLPLHIAAV	2478	4	29
LSPLQALVQ	12	93	PAISLSPQALVQDVC	1889	11	79
LSPLLSITTT	11	79	NSLSPLLSITTTMO	654	7	50
LSPLSITPS	11	71	QMSLSPITPSVHAAV	95	11	79
LSLTLIL	12	86	IVLALSLTLILQIQLH	617	10	71
LIGGFAQH	12	86	DLLIGGFAQHADH	123	12	86
LIVDQIAFL	12	93	FTQIQLDQIAFLSQT	1567	13	93
LTSMALDPS	12	93	YAVLTSMLALDPSIT	2173	9	64
LVAYQATVC	12	86	FPFLVYQATVCADA	1548	9	64
LVBLAGYQ	12	79	GRVBLVLAGYQAVG	1050	9	64
LVGGVLAVL	12	86	FWV1LVGGVLAVLAMY	1664	6	64
LVHLPSVAA	12	100	YKVALVHLPSVAAVLO	1254	12	100
LVNLPLAII	12	79	TEQDVLNLPLAII	1681	10	71
LVVHNDVU	11	79	DLYVLTIVDQDVLV	1134	11	79
LVVQVCAV	12	79	FTALVQVCAVPLV	1094	11	79
LVVLTATP	12	86	QAVLVLATATPOTQ	1305	11	79
LVWDMALM	12	86	APLWDMALM	2069	11	79
LVMPCEZON	12	86	ANALW1QCEZON	2238	12	86
LYTGLQVAN	12	79	ITPLYGLQVANQEVIT	1627	9	64
MANKEEVV	12	86	ITTMANKEEVV	2509	9	64
MAYWQMMWW	12	86	QIYHMMWQMMWW	315	12	86
MCQATIRE	12	86	TCQGQATIRE	2243	12	86
METIPVVA	11	79	ACQKQVVAQVVA	131	11	79
METIPVVA	11	79	LTSMALDPSVAAVLO	2176	6	57
METIPVVA	11	100	YKVALVHLPSVAAVLO	1910	14	100
METIPVVA	12	86	YKVALVHLPSVAAVLO	2793	10	71
METIPVVA	14	100	AKIYMMFSCPQ	1767	12	66
MWCGVHEH	14	100	KHMYWQGQEVILNA	633	5	36
VAQALVAPK	12	86	GRVQVQVVAQVMS	1861	7	50
VARLHAPTO	12	86	TRQVAVLQIPTQSK	1227	6	43
YATDAIMTO	12	86	YVQHATDAIMTO	1437	6	43
YATDAIMTO	14	100	YVQHATDAIMTO	1589	6	43
YAYDAYVCA	12	86	PIVYVATYVCA	10	71	79
YCAVLMHII	12	86	TEATVTTSPDPP	1899	10	71
YCEKMLYD	14	100	VEVYCAVUAVQD	2619	11	71
YDDCLKEPV	12	86	CHYTCERKAVYDVS	1552	6	43
YCTITVYNA	11	79	GLPQCGQLEFRESV	1196	11	79
YCFDPEKQ	12	86	FLAVCTTQVAVAVDF	2524	10	71
YFTDNPSP	12	86	IKMFDQDFPDKR	1211	10	71
YFTLTHID	12	79	TSFPTDNPSPAVP	1563	6	43
YGVVLAAL	12	86	WESFTGVLTHIDAF	1865	12	86
YGVVLAAL	12	93	WVLYGVVLAALAAVC	28	13	93
YVSQUCPEP	12	86	CONGOMYLRTCP	2158	6	43
YVIVVCAAL	12	79	CTVVSQUCPEP	1898	11	79
YVDCNIVCV	12	86	ALWVQVCAAL	1433	12	86
YVDTTCQF	11	79	FDSVNDTTCQF	119	11	79

HCV DRs-Sucre Motif Binding Data Not Included

Core Sequence	Core Freq	Core Conservancy (%)	Exemplary Sequence	Position In HCV Poly/Protein	Exemplary Sequence Frequency	Exemplary Sequence Conservancy (%)
VIAALAYC	12	86	VAGVAVLAVCLIT	1668	8	57
VIAATATPPG	13	93	RLVVLATATPPGSVT	1347	9	64
VEEDGIVNTA	12	86	GVVVEDGIVNTATV	154	12	66
VINPISVIAAT	14	100	KVLVNPIVSIAATGOF	1255	14	100
VITSNALTOP	13	93	DVALITSNALTOPSHI	2172	9	64
VIITSGCNT	11	79	ASGVLITSGCNTTC	2734	10	71
VIVDILAGY	11	79	LGKIVVIVDILAGYDAG	1849	10	71
VIVGGVLLAA	12	86	STIVVLLGGVLLAA	1663	12	86
VIVLNPSSVA	14	100	QYKVLVLPNSVAAATL	1253	14	100
VINLPALIS	12	86	EDLYNLIPALISPGTA	1882	11	79
VPESDAAR	12	86	THYDPESDAAARVTO	1937	7	50
VITSNALVG	12	86	LEVITSNALVGVL	1658	12	86
VIVDALMT	11	79	DIVVIVDALMTGIVT	1436	6	43
VICAILLR	11	79	WVIVGCAALLRING	1899	10	71
VIGVCAAI	11	79	QALVIGVCAAILR	1895	11	79
VILVATATPP	12	86	ARLVVATATPPGSV	1346	9	64
VICFTPSPV	13	93	QCPYVCFTPSPMV	506	13	93
WAGWLLSPR	12	86	GGCIVWGLLSPRSPR	90	5	36
WARMILMTH	12	86	PTLWARMILMTHFS	2870	11	79
WGADTAACG	12	86	ITIWGADTAACGDI	988	6	43
WGPTOPPAR	12	86	FPSWAGPTOPPARSEN	104	10	71
WMHFLIATA	14	100	AVHWMHFLIATAFSG	1917	14	100
WILLAPITA	11	79	SKGIVLLAPITAYAQ	1025	4	29
WITGALIPC	11	79	SYTWTGALIPCIAAE	9	11	64
WYELTPAET	12	86	QCACWYELTPAETVNR	1525	5	36
YATGNAPOC	12	86	QWNTAGNAPOCFS	161	11	79
YCTPSPVN	13	93	GPVNCVTPSPWAVT	507	13	93
YDQVWVHWE	11	79	CHCYDQVWVHWEVTL	1523	10	71
YHODC	12	86	CGYDQVWVHWEVTL	1312	10	71
YIDELITSC	13	93	OPEDOLELITSCSN	2806	11	79
YIAGVAGAL	12	86	UGYAGVAGAL	1857	11	79
YIOTPOO	11	79	CSSTGIVTOSCPRE	2641	10	71
YIPLADCG	11	79	YSTYRPLACDCCS3	1298	10	71
YKVULNPS	14	100	AUGYKVULNPSVAA	1251	11	79
YLAQSLTP	14	100	GQYLAQSLTPVNP	1776	14	100
YQGSSQCP	12	86	PVSYTGQGSSQCP	1162	6	43
YLTROPTIP	11	79	RYMLTROPTIPALAR	2833	9	64
YOAIVCARA	13	93	LVATOAIVCARAOAP	1591	11	79
YRGUDIVSVI	14	100	YAYFGRDUDIVSVI	1420	7	50
YRQAVONE	11	79	PILYRQAVONEVTL	1628	9	64
YRQAVONE	13	93	NCQYRQAVONEVTL	2726	10	71
YRQAVONE	11	79	QACYSEPRDUDONI	2802	6	43
YSGCENPR	11	79	LHSYSGCENPRMASC	2927	6	57
YVGDGGSV	12	86	SAVYVGDGGSVFLY	273	6	57
YVGYLLPNR	11	79		3036		

Table XIXb. IICV DR Surface Multif With Binding Data

Core Sequence	Exemplary Sequence	DR1	DR2+2.1	DR3+2.2	DR3	DR4+4	DR4+5	DR5+11	DR5+12	DR6+2	DR6+9	DR7	DR8	DR9	DR+5.1
FGAYASVQI	TGIGDGVSKVQDQ	0.0150	0.0120	0.0013		0.4200	0.0250	0.0210		0.0001	0.0035	0.0250	0.0050	0.0270	
FGTYWNSI	GAAGFGCTHAASTGFT	0.0190				0.0005									
FIKQALCLL	AECHKQALQLOIA														
FLALLSCL	FSFLALLSCLTVIP														
FPOLOMTC	UNFPOLQGVTCENDA														
FOVAKLHP	POUFOVAAHAPTOGS	0.2100					0.0053								
FRAAVCTD	YQIFRAAVCTQVAK														-0.0003
FSFLALL	GCSEFSFLALLSCL	0.0060					0.0015								0.0030
FSLQDFTF	TYDFSLQDFIETT	0.0001					0.1500								0.0005
FTEAMTRYS	UNFTEAMTRYSAPP														
FTPSWWQ	YVCFTPSPWVYTTD														
FTTLPALST	PCSFITLPALSTGJ	0.0160	-0.0031	-0.0003		0.0920	0.0570	0.0056		-0.0001	0.0035	0.0740	0.1100		
FWAHARHF	LEFHWAHARHMFQ														
FDNFSLQI	LTUDNAFLSOKHOA														
IGKTCVTO	DSWDCNCHVTOV														
IDLTCGFA	GRMLDTQFADLM														
IEARLWHD	AQEAEMLAWHEDQ														
IFLLALSC	SESEFLALLSCLTY														
IGLQWVQD	UFNGLQWVANQDAP														
IGLQWVAD	STILGIGTMDQAE														
ILFRNGFO	CAMLRPAHDFRGED														
ILSPQALW	LPALSPQALWVQVVA														
INAYTTC	TPFRMATTGTOPCPS														
IPVQAPQD	MCYTFVQAPDQAA														
ITTNESEK	QGQNTTVESEKQV														
ITSCSAAVS	LEU15CSAAVSVAH	0.0245	0.0200	-0.0003		0.0070	0.0150	0.0008		0.0510	-0.0003	0.0350	0.0130		
ITPQDLSV	ANPQDPOVQVNE	0.0053				0.0017						0.0004			
IAALAYCL	QDVLVALVAYCLQ														
IAQGCGQ	QHADQCGCGQ														
IAQLSLSD	IQQLSLSDIPOINTA														
IAQTGQV	VIQILQYQVQVQVQV														
IAZTIPQS	LVLVLTATPPQSIV														
ILOPTFET	DFSLQDIFTETTV														
ILQNEATGKA	QVVLQDQAEAGTANLV														
ILEUTSCESS	EVOLEUTSCESSY														
ILEVNTSTWV	SADLEEVNTSTWVQD														
ILFLILLADAT	WFLFLILLADATRVC3														
ILGIVWVQD	FNFLGIVWVQDAPP														
ILQYDFTD	TTIGLQYDFTDNEF														
ILQRATRKT	GPRLOVQVATRKTSER														
ILQRCTERKA	FROLGIVVYKEWALY														
ILQSAFSL	IENQYQSAFSLH	0.0001													
ILQPTFLY	KPTLQGPTFLYFLY	0.0000													-0.0002
ILQHIVVQD	LILQHIVVQDQVLY														
ILSYSPCEI	AFSUSYSPCEI	0.0042													0.0035
ILAFASDN	MATLUFASDNQWV	0.0160	1.9000	0.0130	0.0068	0.0003	0.0079	0.0003	0.0010	0.4400	0.0210	0.4600	0.1100	0.0224	
ILANKLWR	DAQLEAANLWVREDA	0.0008													-0.0005
ILFCHSIVK	GRFLUFLDQFQDCE	0.0001													
ILTCSCEAV	QELUTSCESSAVTA														
ILALLSCL	SPFLALLSCLTVPA														
ILFLILLADA	YVFLFLILLADARVC														
ILFLNQGKW	QFLFLNQGKWWVNA														
ILLADAVC	LFLFLILLADAVC														
ILALSPQD	LFLFLILLADAVC														
ILAQYPLVQD	FADLQDQYPLVQAPL														

Policy for Submitting Manuscripts

Lucy de Smet. Meir will tell Data

UCY DIR Super Muir With Blinding Data

Table XXXb
AlCY DIR 3A Motif With Blinding Information

Core Sequence	Exemplary Sequence	DR3	DR1	DR2w211	DR2w212	DR4w4	DR4w15	DR5w11	DR5w12	DR5w19	DR7	DR8w2	DR9	DR9w2	DR9w3
FLGGCSD	YGFGLDGCGSGAY														
FSLOPFTI	TVDFSLDPTFETIETT														
LEGEPCDQ	MPLLEGERGPQDQSD	-0.0017	0.0001												
LPCEPEPV	GSQLPCPEPQDQVNL	-0.0017													
MADIDAKRM	GHBRMWDMMAMRMAMSP														
MLTQPSHT	LTSALTOPSHIATET														
MSADELWYT	MACMSADLEVYFTSTW														
VATDLMITQ	VVVAVIDALMITYTQ	1.1000	0.0040	0.0047	0.0014										
VCQDHLEFW	GLPVQCDHLEFWESV														
VFPDLGIVV	FLWPPDLOVRYCEK														
VFTDNSSPP	RSPVFTDNSSPPAVP														
VLCCECYDQ	DSVILCECYDQAGAW														
VLEDGVNYA	GYNVNLLEDGVNYATGN														
VLDILAQY	LGKVLVLDLAGYAQ														
VQEPGGER	VFCVQEPKEGTRPKNN														
YDLELTSC	QPEYDLELTSCSSSN														
YSIPLDLP	GACYSIPLDLP														
YVGOLCGSVFLV	SAWYVGOLCGSVFLV														
YVPESDAAA	PTIYVPESDAAAART														

Table XXc UCYJII Motif

Core Sequence	Core Freq.	Core Conservancy (%)	Exemplary Sequence	Position In HCV Polyprotein	Exemplary Sequence Frequency	Exemplary Sequence Conservancy (%)
FGISKKQD	14	100	I M F C H S K K Q D K D E A	1395	14	100
FSTPQFD	11	79	P F S T P Q F D P Q F S T V	2687	11	79
LAEGPKDA	12	88	G H Q L A E G P K D A G Q G	1726	6	57
URPLIAQPT	11	79	U R P L I A Q P T L	1616	10	71
VRATRTESE	11	79	F L G R T R T E S E Q	43	10	71
YLTRHADV	12	86	S D L Y L V I T H A D V I P V	1133	11	79
ASTHPAPCR	11	79				

Table XXd HCV 3B Motif Binding Data

Core Sequence	Exemplary Sequence	DR1	DR2w11	DR2w212	DR3	DR4w4	DR4w15	DR5w11	DR5w12	DR6w10	DR6w2	DR7	DR8	DR9w43
FO-6886CD FSTDTTPO VADPKRKA LKPFLGPT VATRATSE YLVTRHADY KSTKTPOR	HUCFISRSKQKDELA PAPFSTDTTPOSTV QVQLEQGPKRKAQG LURKFLGPTL FLGPAATATSESGO SDLYLVTRHADYIPV													0.0180
														0.0022

TABLE XXI. Population coverage with combined HLA Supertypes

<u>HLA-SUPERTYPES</u>	<u>PHENOTYPIC FREQUENCY</u>					Average
	Caucasian	North American Black	Japanese	Chinese	Hispanic	
<u>a. Individual Supertypes</u>						
A2	45.8	39.0	42.4	45.9	43.0	43.2
A3	37.5	42.1	45.8	52.7	43.1	44.2
B7	38.6	52.7	48.8	35.5	47.1	44.7
A1	47.1	16.1	21.8	14.7	26.3	25.2
A24	23.9	38.9	58.6	40.1	38.3	40.0
B44	43.0	21.2	42.9	39.1	39.0	37.0
B27	28.4	26.1	13.3	13.9	35.3	23.4
B62	12.6	4.8	36.5	25.4	11.1	18.1
B58	10.0	25.1	1.6	9.0	5.9	10.3
<u>b. Combined Supertypes</u>						
A2, A3, B7	83.0	86.1	87.5	88.4	86.3	86.2
A2, A3, B7, A24, B44, A1	99.5	98.1	100.0	99.5	99.4	99.3
A2, A3, B7, A24, B44, A1, B27, B62, B58	99.9	99.6	100.0	99.8	99.9	99.8

SF 184895 v1

Table XXII
ICCV ANALOGS

Sequence	AA	Fixed Name.	Motif	A1	A2	A3	A4	B7	1*
	9	IVXEKMALY							Anchor Fixer
	9	AVXTRGVAK							
	9	EVFXNQPEK							
	9	HIFXHSKK							
	9	LPQXSFSIF							
	9	LFJFHSKKK							
	10	VIAALAAYXL							
	10	HIFXHSKKK							
	10	AXXNWTRGERA							
	10	YLLPRGRGPPV							
	9	FPQCSFSIF							
	9	LPVCSFSIF							
	9	LPQCSFSYF							
	9	LPQCMFSIF							
	9	LPFCFSIF							
	9	LPQCSFSFP							
	9	LPQCSFSII							
	9	PPVWHGCP							
	10	KPTLHGPTPI							
	10	APTLWARMII							
	9	SPAGSAPSII							
	10	UPAGFALGI							
	9	SPGGRVIFI							
	9	LPQCSFSII							
	9	DPRNINSHII							
	9	SPGALVVGVI							
	10	TPLLYYALGAI							
	9	TSQGVVLWQV							
	9	SIISGVVLWQV							
	9	SLMAFTASV							
	9	GLRQDCTMVL							
	10	KLYALGVNAY							
	10	YLLPSRGPKL							
	10	KLSGALGLNAV							
	10	YLPFRGRGPAL							
	10	VFFMILLGGWV							
	9	KVLSGLGVNAV							
	9	CINGVCMTA							
	9	CANGVCMTA							
	10	LV2.L10							
	10	LV2.VA9							
	10	IA2.V9							

LCY ANALOGS

AA	Sequance	Fixed Nomen.	A ¹ Molli	A ² Super Molli	A ³ Super Molli	A ²⁴ Molli	B ⁷ Super Molli	1 [*] Anchor Fixer
9	CYNGVCAV			N	Y	N	N	

40

Table XXXII. Immunogenicity of identified supermotif-bearing peptides

Supermotif	Peptide	Sequence	Protein	Position	Immunogenicity						Frequency	Response
					Barnaba; patients	Barnaba; contacts	Human ^a	Chisari	Pape	overall		
A2	1073.05	LLFNILGGWV	NS4	1812	1/6	7/17	2/21	0/6	10/50	6/6	6.4 (1.7)	
	1090.18	FLLLADARV	NS1/E2	728	2/6	7/17	1/21	0/6	10/50	5/6	9.5 (3.0)	
	1013.02	YLVAYQATV	NS4	1590	1/6	4/17	1/21	0/6	6/50	5/6	8.5 (3.7)	
	1090.22	RLIVFPDLGV	NS5	2578	2/6	5/17	0/21	0/6	7/50	0/6		
	1013.1002	DLMGYIPLV	Core	132	2/6	7/17	1/21	1/6	11/50	5/6	8.8 (2.6)	
	24.0073	WMNRJIAFA	NS4	1920	1/6	3/17	2/21	1/6	7/50	0/6		
	24.0075	VLVGGVIAA	NS4	1666	1/6	6/17	3/21	1/6	11/50	0/6		
	1174.08	HMWNFISGI	NS4	1769	3/6	3/17	2/21	0/6	8/50	6/6	6.4 (1.7)	
	1073.06	ILAGYGAGV	NS4	1851	2/6	3/17	0/21	0/6	5/50	3/6	54.7 (3.3)	
	1073.07	YLLPRRGPRL	CORE	35	2/6	5/17	7/21	1/6	17/50	4/6	59.1 (7.2)	
	24.0071	LLFLLLADA	NS1/E2	726	2/6	9/17	0/21	0/6	11/50	0/6		
	1.0119	YLVTRHADV	NS3	1131	6/6	10/17	0/21	1/6	17/50	0/6		
A3	1.0952	KTSERSQPR	CORE	51	2/16	1/4	3/12	0/6	6/38	3/6	23.4 (1.3)	
	1073.11	RLGVRA TRK	CORE	43	4/16	1/4	7/12	1/6	13/38	3/6	42.2 (1.2)	
	1.0955	QLFTFSPPR	ENV	290	1/16	0/4	6/12	1/6	8/38			
	1073.13	RMYVGGVEHR	NS1/E2	632	5/16	1/4	4/12	1/6	11/38	2/6	2.8 (1.1)	
	1.0123	LIFCHSKKK	NS3	1396	6/16	1/4	4/12	2/6	13/38	3/6	4.4 (1.1)	
	1073.10	GVAGALVAFK	NS4	1863	3/16	0/4	6/12	2/6	11/38	6/6	56.5 (1.7)	
	24.0090	VAGALVAFK	NS4	1864	4/16	1/4	6/12	0/4	11/38	1/6	7.1	
	24.0086	TLGFGAYMSK	NS3	1262	6/16		2/12	2/5	10/33			
B7	1145.12	LPGCSFSIF	CORE	169			2	3/10	5			

Table XXIV. Human and murine MHC-peptide binding assays established using purified MHC molecules and gel filtration chromatography

A. Class I binding assays

Species	Antigen	Allele	Cell line	Radiolabeled peptide		Notes
				Source	Sequence	
Human	A1	A*0101	Steinlin	Hu. J chain 102-110	YTAVPLVY	no NEN in PI cocktail
	A2	A*0201	JY	HBVc 18-27 F6->Y	FLPSDYPPSV	"
	A2	A*0202	P815 (transfected)	HBVc 18-27 F6->Y	FLPSDYPPSV	"
	A2	A*0203	FUN	HBVc 18-27 F6->Y	FLPSDYPPSV	"
	A2	A*0206	CLA	HBVc 18-27 F6->Y	FLPSDYPPSV	"
	A2	A*0207	721.221 (transfected)	HBVc 18-27 F6->Y	FLPSDYPPSV	"
	A3		GM3107	non-natural (A3CON1)	KVFPYALINK	"
A11			BVR	non-natural (A3CON1)	KVFPYALINK	"
A24		A*2402	KAS116	non-natural (A24CON1)	AYIDNYNKF	"
A31		A*3101	SPACH	non-natural (A3CON1)	KVFPYALINK	"
A33		A*3301	LWAGS	non-natural (A3CON1)	KVFPYALINK	"
A28/68		A*6801	CIR	HBVc 141-151 T7->Y	STLPETYYVRR	"
A28/68		A*6802	AMAI	HBV pol 646-654 C4->A	FTQAGYPAL	"
B7		B*0702		GM3107	APRTLVYLL	"
B8		B*0801		A2 signal seq. 5-13 (L7->Y)	FLKDYQLL	"
B27		B*2705	Steinlin	IVgp 586-593 Y1->F, Q5->	FRYNGLJHR	"
B35		B*3501	LG2	R 60s	PPFKYAAAF	"
B35		B*3502	CIR, BVR	non-natural (B35CON2)	PPFKYAAAF	"
B35		B*3503	TISI	non-natural (B35CON2)	PPFKYAAAF	"
B44		B*4403	EHM	non-natural (B35CON2)	AEMGKYSFY	"
B51			PITOUT	EF-1 G6->Y	PPFKYAAAF	"
B53		B*5301	KAS116	non-natural (B35CON2)	PPFKYAAAF	"
B54		B*5401	AMAI	non-natural (B35CON2)	PPFKYAAAF	"
Cw4		Cw*0401	KT3	non-natural (C4CON1)	QYDDAVYKL	"
Cw6		Cw*0602	CIR	non-natural (C6CON1)	YRHDDGNVL	"
Cw7		Cw*0702	721.221 transfected	non-natural (C7CON1)	YRHDDGNVL	"
Mouse	D ^b		EL4	Adenovirus E1A P7->Y	SGPSNTYPEI	"
	K ^b		EL4	VSV NP 52-59	RGYYVQGL	"
	D ^d		P815	HTV-IIIB ENV G4->Y	RGPYRAFVTI	"
	K ^d		P815	non-natural (KdCON1)	KFNPMTKYI	"
	L ^d		P815	HBVs 28-39	IPQLSDSYWTSL	"

Table XXIV. Human and murine MHC-peptide binding assays established using purified MHC molecules and gel filtration chromatography

B. Class II binding assays			Cell line	Radiolabeled peptide		Notes
Species	Antigen	Allele		Source	Sequence	
Human	DR1	DRB1*0101	LG2	HA Y307-319	YPKYVKONTLKLAT	
	DR2	DRB1*1501	L466.1	MBP 88-102Y	VVHFKNIVTPRIPPY	
	DR2	DRB1*1601	L242.5	non-natural (760.16)	YAAFAAAKTAAGAAFA	
	DR3	DRB1*0301	MAT	MT 65kD Y3-13	YKTAFFDEEARR	optimal assay pH is 4.5
	DR4w4	DRB1*0401	Preiss	non-natural (717.01)	YARFQSQTTLKQKT	
	DR4w10	DRB1*0402	YAR	non-natural (717.10)	YARFQRQTTLKAAGA	
	DR4w14	DRB1*0404	BIN 40	non-natural (717.01)	YARFQSQTTLKQKT	
	DR4w15	DRB1*0405	KT3	non-natural (717.01)	YARFQSQTTLKQKT	
	DR7	DRB1*0701	Pitout	Tet. tox. 830-843	QYIKANSKFIGITE	
	DR8	DRB1*0802	OLL	Tet. tox. 830-843	QYIKANSKFIGITE	
	DR8	DRB1*0803	LUY	Tet. tox. 830-843	QYIKANSKFIGITE	
	DR9	DRB1*0901	HID	Tet. tox. 830-843	QYIKANSKFIGITE	
	DR11	DRB1*1101	Sweig	Tet. tox. 830-843	QYIKANSKFIGITE	
	DR12	DRB1*1201	Herluf	unknown eluted peptide	EALIHQLKINPYVLS	
	DR13	DRB1*1302	H0301	Tet. tox. 830-843 S->A	QYIKANAKFIGITE	
	DR51	DRB5*0101	GM3107 or L416.3	Tet. tox. 830-843	QYIKANAKFIGITE	
	DR51	DRB5*0201	L255.1	HA 307-319	PKYVKQNTLKLAT	
	DR52	DRB3*0101	MAT	Tet. tox. 1272-1284	NGQIQLNDPNRDIL	
	DR53	DRB4*0101	L257.6	non-natural (717.01)	YARFQSQTTLKQKT	no NEM in PI mix
	DQ3.1	DQA1*0301/DQB1*0301	PF	non-natural (ROI ^Y)	YAHAAAHAAAHAAAHAA	
Mouse	IA ^b		DB27.4	non-natural (ROI ^Y)	YAHAAAHAAAHAAAHAA	optimal assay pH is 5.5
	IA ^d		A20	non-natural (ROI ^Y)	YAHAAAHAAAHAAAHAA	
	IA ^c		CH-12	HEL 46-61	YNTDGSTDYQILQNSR	optimal assay pH is 5.0
	IA ^c		LS102.9	non-natural (ROI ^Y)	YAHAAAHAAAHAAAHAA	
	IA ^c		91.7	non-natural (ROI ^Y)	YAHAAAHAAAHAAAHAA	
	IE ^d		A20	Lambda repressor 12-26	YLEDARRKKAIEYKKK	optimal assay pH is 5.0
	IE ^c		CH-12	Lambda repressor 12-26	YLEDARRKKAIEYKKK	optimal assay pH is 5.0

Table XXXV. Monoclonal antibodies used in MHC purification.

Monoclonal antibody	Specificity
W6/32	HLA-class I
B123.2	HLA-B and C
IVD12	HLA-DQ
LB3.1	HLA-DR
M1/42	H-2 class I
28-14-8S	H-2 D ^b and L ^d
34-5-8S	H-2 D ^d
B8-24-3	H-2 K ^b
SF1-1.1.1	H-2 K ^d
Y-3	H-2 K ^b
10.3.6	H-2 IA ^k
14.4.4	H-2 IE ^d , IE ^K
MKD6	H-2 IA ^d
Y3JP	H-2 IA ^b , IA ^s , IA ^u

Table XXVI: HCV-derived conserved high algorithm A*0201-binding peptides

Peptide	Molecule	1st Position	Sequence	Cons.	A2-supertype binding capacity (IC50 nM)				
					A*0201	A*0202	A*0203	A*0206	A*6802
1073.05	NS4	1812	LLFNILGGWW	85	4.2	113	3.2	19	33
1090.18	NS1/E2	728	FLLADARV	92	18	90	149	247	111
1013.02	NS4	1590	YLVAYQATV	85	20	39	16	82	33
1090.22	NSS	2611	RLIVFPDLGV	79	56	391	10	370	8000
1013.1002	CORE	132	DLMGYIPLV	79	80	4778	204	481	12
24.0073	NS4	1920	WMNRLIAFA	100	122	130	3.3	1609	400
24.0075	NS4	1666	VLVGGVLAA	85	185	331	32	308	3077
1174.08	NS4	1769	HMWNFISGI	92	15	10750	77	132	7547
1073.06	NS4	1851	ILAGYYGAGV	79	116	143	5.0	755	889
1073.07	CORE	35	YLLPRRGPRL	92	125	6143	455	416	10256
24.0071	NS1/E2	726	LLFLFLLADA	100	217	287	455	3364	3077
1.0119	LORF	1131	YLVTRHADV	85	455	2048	3.6	71	3077
24.0065	NS4	1891	ILSPGALVV	92	238	10750	27	1028	3077
1013.12	NS1/E2	686	ALSTGGLIHL	85	313	7167	45	18500	10256
939.14	NS1/E2	696	HLHQNIVDV	85	500	3071	19	1370	10811
1090.21	NSS	2918	RLHGIGSAFSL	79	179	782	625	18500	12500

Table XXVII: HCV-derived conserved high algorithm A*03 and/or A*11 binding peptides

Peptide	Molecule	1st Position	Sequence	Cons.	A3-supertype binding capacity (IC50 nM)				
					A*03	A*11	A*3101	A*3301	A*6801
1.0952	CORE	51	KTSERSQPR	92	69	94	67	1813	145
1073.11	CORE	43	RLGYRATRK	79	12	207	429	-	3
1.0955	ENV1	290	QLFTFSPPR	79	15	182	621	3766	3
1073.13	NS1/E2	632	RMYVGGVEHR	100	15	300	95	9667	1778
1.0123	NS3	1396	LIFCHSKKK	100	20	32	2535	24167	333
1073.10	NS4	1863	GVAGALVAFK	85	28	4	3273	26364	118
24.0090	NS4	1864	VAGALVAFK	85	46	7	3750	11600	258
24.0086	NS3	1262	LGFGAYMSK	85	136	21	2950	22308	222
1174.16	NS1/E2	557	WMNSTGFTK	79	208	74	12857	690	1429
1073.14	NS3	1261	TLGFGAYMSK	85	136	98	-	22308	8889
1090.23	LORF	1183	AVCTRGVAK	79	423	240	16364	-	2
1090.24	NSS	2596	EVFCVQPEK	85	13750	222	-	-	2
24.0103	NS1/E2	647	AACNWTRGER	85	36667	429	400	5273	4444
1073.16	NS3	1232	HLHAPTGSGK	85	19	2500	-	-	2857
1073.12	NS3	1395	HLIFCHSKKK	100	423	-	20000	-	1
1090.26	NS3	1395	HLIFCHSKK	100	440	10000	-	8000	1

* A dash indicates IC50nM >30,000

Table XXXVIII: HCV derived conserved B*0702 binding peptides

Peptide	Molecule	1st Position	Sequence	Conserv.	B7-supertype binding capacity (IC50 nM)				
					B*0702	B*33501	B*51	B*5301	B*5401
1145.12	Core	169	LPGCSFSIF	92	28	90	100	114	6667
15.0048	E2	681	LPALSTGLI	85	157	-	2.8	1500	20000
15.0234	NS3	1620	KPTLHGPTPL	79	3.9	-	27500	-	1
15.0247	NS5	2835	APTLWARMIL	79	6.3	-	5500	-	1
15.0042	CORE	99	SPRGSRPSW	79	14	-	11000	-	1
15.0039	Core	57	QPRGRRQPI	92	24	-	-	-	1
15.0218	Core	37	LPRRGPRGLGV	92	29	-	6111	-	4000
15.0060	NS5	2615	SPGQRVFEFL	79	46	-	27500	-	1
15.0043	Core	111	DPRRRSRNL	85	324	-	-	-	1
15.0063	NS5	2835	APTLWARMI	79	344	-	4583	-	1
1292.17	NS5	2317	PPVVGCP	79	393	-	-	-	1
15.0239	NS4	1893	SPGALVYGVV	79	423	-	3438	-	1
15.0235	NS3	1621	TPLLYRLGAV	92	458	-	6875	-	909

Table XXVIII: HCV derived conserved B*0702 binding peptides**B. Additional HCV derived B7 supermotif peptides.**

Peptide	Molecule	1st Position	Sequence	Consy.	B7-superotype binding capacity (IC50 nM)				
					B*0702	B*3501	B*51	B*5301	B*5401
29.0035	NS3	1378	IPFYGKAI	92	458	-	46	-	50
29.0040	Core	37	LPRRGPRL	92	0.85	-	306	-	5000
29.0036	Core	137	IPLVGAPL	79	13	2250	79	-	2857
16.0187	NS1/E2	680	LPCSFITLPA	64	423	24000	9167	-	15
29.0039	Core	169	LPGCSFSI	92	500	200	932	620	6250
15.0219	Core	142	APLGGAARAL	71	9.5	-	-	-	12500
29.0031	NS5	2869	APTLWARM	79	13	-	4583	-	4348
15.0231	NS3	1512	RPSCGMFDSSV	71	153	-	-	-	1
29.0085	NS5	2474	LPINALSNSL	57	220	18000	1170	-	11111
29.0037	NS5	2608	KPARLIVF	85	367	-	3235	-	16667
15.0237	NS4	1789	NPAIASLMAF	71	393	9000	5000	-	1
29.0118	NS5	2869	APTLWARMILM	79	423	-	-	3030	1
29.0042	NS4	1720	LPYIEQGM	85	423	-	1375	-	7692

C. Engineered analogs of B7 supermotif peptides.

Peptide	Molecule	1st Position	Sequence	Consy.	B7-superotype binding capacity (IC50 nM)				
					B*0702	B*3501	B*51	B*5301	B*5401
1145.12	Core	169	LPGCSFSIF	92	28	90	100	114	6667
1292.24	Core	169	LPGCSFSII	37	4364	5.3	262	1056	3
1145.13	Core	169	FPGCSFSIF	19	1.6	132	3.2	6.7	5

* A dash indicates IC50 nM >30,000.

Table XXIX: HCV-derived A1- and A24-motif containing peptides**A. A1-motif peptides**

Peptide	Molecule	Position	Sequence	Conserv.	HLA-A*0101 binding (IC50 nM)
13.0019	NS5	2922	LSAFSLHSY	79	31
1.0509	NS5	2921	GLSAFSLHSY	79	61
1069.62	NS3	1128	CTCGSSSDLY	79	68
24.0093	NS5	2129	EVDGVRLHRY	100	167
13.0016	NS3	1241	KSTKVPAAY	85	1923
1.0125	NS3	1525	CYDAGCAWY	79	4032
24.0008	E1	206	DCSNSSIVY	85	16667
24.0094	NS5	2720	TNSKGQNCGY	100	-
24.0096	NS3	1240	GKSTKVPAAY	85	-
24.0100	NS3	1292	TGAPITYSTY	85	-
	NS3	1263	VAATLGFGAY	100	-
	NS5	2639	VMGSSYGFQY	79	-
	NS5	2640	MGSSYGFQY	79	-

A dash indicates IC50 nM >25000

B. A24 -motif peptides

Peptide	Molecule	Position	Sequence	Conserv.	HLA-A*2402 binding (IC50 nM)
24.0092	NS4	1765	FWAKHMWNF	85	1.7
13.0075	NS4	1778	QYLAGLSTL	100	250
1073.18	NS1/E2	636	MYVGGVEHRL	92	444
13.0074	NS3	1297	TYSTYKGFL	85	522
13.0134	NS5	2647	QYSPGQRVEF	79	667
24.0091	NS4	1772	NFISGIQYL	100	706
13.0131	Core	135	GYIPLVGAPL	79	2105
24.0108	Core	173	SFSIFLLALL	100	2927
13.0132	NS3	1248	AYAAQGYKVL	79	13333
13.0133	NS4	1859	GYGAGVAGAL	85	-
1174.08	NS4	1769	HMWNFISGI	93	-
	E1	317	RMAWDMMMNW	85	-
	NS1/E2	635	RMYVGGVEHRL	93	-
	NS3	1422	YYRGLDVSFI	100	-
	NS3	1468	DFSLDPFTFI	100	-
	NS3	1608	SWDQMWKCL	79	-
	NS3	1664	TVVLVGGVL	85	-
	NS4	1732	QFKQKALGL	85	-
	NS4	1732	QFKQKALGLL	85	-
	NS4	1765	FWAKHMWNFI	85	-
	NS4	1919	QWMNRLIAF	100	-
	NS5	2241	LWRQEMGGNI	85	-
	NS5	2669	GFSYDTRCF	79	-
	NS5	2875	RMILMTHFF	85	-

A dash indicates IC50 nM >25000

Table XXX: Immunogenicity of A2-supertype cross-reactive binders

Peptide	Sequence	Protein	Position	Immunogenicity			overall	Frequency	Response
				Human ^a	Barnaba; Barnaba; Chisari	Pape			
1073.05	LLFNILGGWV	NS4	1812	1/6	7/17	2/21	0/6	10/50	6/6 6.4 (1.7)
1090.18	FLLLADARV	NS1/E2	728	2/6	7/17	1/21	0/6	10/50	5/6 9.5 (3.0)
1013.02	YLVAYQATV	NS4	1590	1/6	4/17	1/21	0/6	6/50	5/6 8.5 (3.7)
1090.22	RLIVFPDLGV	NS5	2578	2/6	5/17	0/21	0/6	7/50	0/6
1013.1002	DLMGYIPLV	Core	132	2/6	7/17	1/21	1/6	11/50	5/6 8.8 (2.6)
24.0073	WMNRLLJFA	NS4	1920	1/6	3/17	2/21	1/6	7/50	0/6
24.0075	VLVGGVLA	NS4	1666	1/6	6/17	3/21	1/6	11/50	0/6
1174.08	HMWNFISGI	NS4	1769	3/6	3/17	2/21	0/6	8/50	6/6 6.4 (1.7)
1073.06	ILAGYGAGV	NS4	1851	2/6	3/17	0/21	0/6	5/50	3/6 54.7 (3.3)
1073.07	YLLPRRGPRL	CORE	35	2/6	5/17	7/21	1/6	17/50	4/6 59.1 (7.2)
24.0071	LLFLLLADA	NS1/E2	726	2/6	9/17	0/21	0/6	11/50	0/6
1.0119	YLVTRHADV	NS3	1131	6/6	10/17	0/21	1/6	17/50	0/6

a. Data shown represents the number of positive responses over the total number of patients or contacts examined.

b. Frequency represents the number of positive responses over the total number of mice examined. Response indicates the average magnitude (standard deviation) of the response in positive animals, measured in lytic units.

Table XXXI: Immunogenicity of A3-supertype cross-reactive binders

Peptide	Sequence	Protein	Position	Immunogenicity				Frequency	Response	
				Human ^a		Transgenic mice ^b				
Barnaba; Barnaba;										
1.0952	KTSERSQPR	CORE	51	2/16	1/4	3/12	0/6	6/38	3/6 23.4 (1.3)	
1073.11	RLGVRATRK	CORE	43	4/16	1/4	7/12	1/6	13/38	3/6 42.2 (1.2)	
1.0955	QLFTFSPRR	ENV	290	1/16	0/4	6/12	1/6	8/38		
1073.13	RMYVGGVEHR	NS1/E2	632	5/16	1/4	4/12	1/6	11/38	2/6 2.8 (1.1)	
1.0123	LJFCHSKKK	NS3	1396	6/16	1/4	4/12	2/6	13/38	3/6 4.4 (1.1)	
1073.10	GVAGALVAFK	NS4	1863	3/16	0/4	6/12	2/6	11/38	6/6 56.5 (1.7)	
24.0090	VAGALVAFK	NS4	1864	4/16	1/4	6/12	0/4	11/38	1/6 7.1	
24.0086	TLGFGAYMSK	NS3	1262	6/16	2/12	2/5	10/33			

a. Data shown represents the number of positive responses over the total number of patients or contacts examined.

b. Frequency represents the number of positive responses over the total number of mice examined. Response indicates the average magnitude (standard deviation) of the response in positive animals, measured in lytic units.

Table XXXII. Candidate HCV-derived HTL epitopes

Selection criteria	Peptide	Sequence	Source	Conservancy	
				Total	Core
A. DR-supermotif conserved 15mers	I283.01	GQIVGGVYLLP RRGPR	HCV Core 28	93	93
	I283.02	VYLLP RRGPR LGVRA	HCV Core 34	93	93
	I283.03	GWLLSPRGSRPSWGPT	HCV Core 95	79	79
	I283.04	LGKVIDTLCGFADL	HCV Core 119	79	86
	I283.05	IDTLTCGFADLMGYI	HCV Core 123	86	86
	I283.06	ADLMGYIPLVGAPLG	HCV Core 131	79	79
	I283.07	GVRVLEDGVNYATGN	HCV Core 154	86	86
	I283.08	GVNYATGNLPGCSFS	HCV Core 161	79	86
	I283.09	GCSFSIFLLALLSCL	HCV Core 171	86	100
	I283.10	GHRM AWDMMMNWSPT	HCV E1 315	86	86
	I283.11	CGPVYCFPS P VVVG	HCV NS1/E2 506	93	93
	I283.12	VYCFPS P VVVGTTD	HCV NS1/E2 509	93	93
	I283.13	GNWFGCTWMNSTGFT	HCV NS1/E2 550	79	86
	I283.14	FTTLPA LSTGLIHLH	HCV NS1/E2 684	79	86
	I283.17	DLYL VTRHADVIPVR	HCV NS3 1134	79	79
	I283.18	RAAV CTRGVAKAVDF	HCV NS3 1186	79	79
	I283.20	AQGYKVLV LNP SVA A	HCV NS3 1251	79	100
	I283.21	GYKV LVLNP SVA ATL	HCV NS3 1253	100	100
	I283.22	VLVLNP SVA ATL GFG	HCV NS3 1256	100	100
	I283.23	GTVLDQ AETAGARLV	HCV NS3 1335	86	86
	I283.24	GARL VVLA TATPPGS	HCV NS3 1345	79	86
	I283.25	GRHLIFCHSKKKCDE	HCV NS3 1393	100	100
	I283.27	DSVIDC NTCVTQTV D	HCV NS3 1454	86	86
	I283.28	TVDFSLDPTFTIETT	HCV NS3 1466	79	100
	I283.30	FTGLT HIDA HFLSQ T	HCV NS3 1567	93	93
	I283.31	YL VAYQATV CARA Q A	HCV NS3 1591	79	93
	I283.32	KPTLH GPTPLL YRLG	HCV NS4 1620	79	79
	I283.33	LEVVTSTW VL VGGV L	HCV NS4 1658	86	86
	I283.34	TW VL VGGV LA ALA AY	HCV NS4 1664	86	86
	I283.35	AEQFKQKAL GLLQTA	HCV NS4 1730	86	86
	I283.40	PAILSPGAL VVG VVCA	HCV NS4 1889	79	93
	I283.41	GAL VVG VVCA AIL RR	HCV NS4 1895	79	79
	I283.42	CAAILRRH VGP GEGA	HCV NS4 1903	79	79
	I283.43	AVQWMNRLIA FASRG	HCV NS4 1917	100	100
	I283.44	MNRLIA FASRG NHVS	HCV NS4 1921	86	100
	I283.48	ANLLW RQEMGGNITR	HCV NS5 2238	86	86
	I283.49	RQEMGGNITR VESEN	HCV NS5 2243	86	86
	I283.52	ARLIVFPDLGVRVCE	HCV NS5 2610	79	79
	I283.53	FPDLGVRVCEK M ALY	HCV NS5 2615	79	100
	I283.54	GVRVCEK M ALYDV VS	HCV NS5 2619	79	100
	I283.56	QPEYD LELITSCSSN	HCV NS5 2808	79	93
	I283.57	LELITSCSSNV SVAH	HCV NS5 2813	79	100
	I283.58	PTLWARMILMTHFFS	HCV NS5 2870	79	86
	I283.59	LHGLSAFSLHSYSPG	HCV NS5 2919	79	79
	I283.60	AFSLHSYSPGEINRV	HCV NS5 2924	79	79

Table XXXII. Candidate HCV-derived HTL epitopes

Selection criteria	Peptide	Sequence	Source	Conservancy	
				Total	Core
B. High algorithm conserved core	1283.15	VVLLFLLLADARVCS	HCV NS1/E2 724	29	100
	1283.16	SKGWRLLAPITAYAQ	HCV NS3 1025	29	79
	1283.19	PQTFQVAHLHAPTGS	HCV NS3 1225	43	85
	1283.26	DVVVVATDALMTGYT	HCV NS3 1436	43	79
	1283.29	WESVFTGLTHIDAHF	HCV NS3 1563	43	92
	1283.45	LTSMLTDPSHITAET	HCV NS5 2176	57	100
	1283.46	ASQLSAPSILKATCTT	HCV NS5 2208	50	79
	1283.47	DADLIEANLLWRQEM	HCV NS5 2232	50	85
	1283.50	SYTWTGALITPCAAE	HCV NS5 2456	64	79
	1283.51	TTIMAKNEVFCVQPE	HCV NS5 2589	64	85
	1283.55	GSSYGFQYSPGQRVE	HCV NS5 2641	71	79
	1283.61	ASCLRKLGVPPRLRVW	HCV NS5 2939	50	85
C. Collaborator	F098.03	AAYAAQGYKVLVNPSVAAT	HCV NS3 1242-1261	71	100
	F098.04	GYKVLVLPNSVAATLGFGAY	HCV NS3 1248-1267	100	
	F098.05	GYKVLVLPNSVAAT	HCV NS3 1248-1261	100	
	F134.01	RRPQDVKFPGGGQIVGGVY	HCV Core 17-35	86	
	F134.02	DVKFPGGGQIVGGVYLLP	HCV Core 21-40	86	
	F134.03	GYKVLVLPNSVAATLGFGAY	HCV NS3 1253-1272	100	
	F134.04	TLHGPTPLLYRQLGAVQNEIT	HCV NS4 1622-1641		79
	F134.05	NFISGIQYLAGLSTLPGNPA	HCV NS4 1772-1791	100	
	F134.06	LLFNILGGWVAAQLAAPGAA	HCV NS4 1812-1831		86
	F134.07	GPGEAVQWMNRLLIAFASRG	HCV NS4 1912-1931	86	100
	F134.08	GEAGAVQWMNRLLIAFASRG	HCV NS4 1914-1934	100	
	Pape 21	AIPLEVKGGRHLIFCHSKR	HCV NS3 1379-1398	21	100
D. DR3 motif	Pape 22	GRHLIFCHSKRKCDELATKL	HCV NS3 1388-1407		100
	Pape 29	SVDCNTCVTQTVDFSLDPT	HCV NS3 1450-1469	86	
	35.0102	GVRVLEDGVNYATGN	HCV 154	86	86
	35.0103	SAMYVGDLCGSVFLV	HCV 273	57	86
	35.0104	GHRMAWDMMMMNWSPT	HCV 315	86	86
	35.0105	SDLYLVTRHADVIPV	HCV 1133	79	86
	35.0106	VVVVATDALMTGYTG	HCV 1437	42	86
	35.0107	TVDFSLDPTFTIETT	HCV 1466	79	100
	35.0108	DSSVLCECYDAGCAW	HCV 1518	71	93
	35.0109	GLPVCQDHLEFWESV	HCV 1552	42	86
	35.0110	GMQLAEQFKQKALGL	HCV 1726	57	86
	35.0111	PTHYVPESDAAARVT	HCV 1936	86	86
	35.0112	GSQLPCEPEPDVAVL	HCV 2162	64	86
	35.0113	LTSMLTDPSHITAET	HCV 2176	57	100
	35.0114	MPPLEGEPGDPDLSD	HCV 2401	79	100
	35.0115	QPEYDLELITSCSSN	HCV 2808	79	93
	1283.25	GRHLIFCHSKKKCDE	HCV NS3 1393-1407		

Table XXXIII. HLA-DR screening panels

Screening Panel	Antigen	Alleles	Representative Assay		Phenotypic Frequencies					
			Allele	Alias	Cauc.	Blk.	Jpn.	Chn.	Hisp.	Avg.
Primary	DR1	DRB1*0101-03	DRB1*0101	(DR1)	18.5	8.4	10.7	4.5	10.1	10.4
	DR4	DRB1*0401-12	DRB1*0401	(DR4w4)	23.6	6.1	40.4	21.9	29.8	24.4
	DR7	DRB1*0701-02	DRB1*0701	(DR7)	26.2	11.1	1.0	15.0	16.6	14.0
	Panel total				59.6	24.5	49.3	38.7	51.1	44.6
Secondary	DR2	DRB1*1501-03	DRB1*1501	(DR2w2 β1)	19.9	14.8	30.9	22.0	15.0	20.5
	DR2	DRB5*0101	DRB5*0101	(DR2w2 β2)	-	-	-	-	-	-
	DR9	DRB1*0901-11,0901-12	DRB1*0901	(DR9)	3.6	4.7	24.5	19.9	6.7	11.9
	DR13	DRB1*1301-06	DRB1*1302	(DR6w19)	21.7	16.5	14.6	12.2	10.5	15.1
	Panel total				42.0	33.9	61.0	48.9	30.5	43.2
Tertiary	DR4	DRB1*0405	DRB1*0405	(DR4w15)	-	-	-	-	-	-
	DR8	DRB1*0801-5	DRB1*0802	(DR8w2)	5.5	10.9	25.0	10.7	23.3	15.1
	DR11	DRB1*1101-05	DRB1*1101	(DR5w11)	17.0	18.0	4.9	19.4	18.1	15.5
	Panel total				22.0	27.8	29.2	29.0	39.0	29.4
Quaternary	DR3	DRB1*0301-2	DRB1*0301	(DR3w17)	17.7	19.5	0.4	7.3	14.4	11.9
	DR12	DRB1*1201-02	DRB1*1201	(DR5w12)	2.8	5.5	13.1	17.6	5.7	8.9
	Panel total				20.2	24.4	13.5	24.2	19.7	20.4

Table XXXIV. HLA-DR binding capacity of target derived peptides: DR-supermotif and algorithm positive peptides.

Peptide	Sequence	Source	Binding capacity (IC50 nM)									DR alleles bound	
			DR1	DR2w2B1	DR2w2B2	DR4w4	DR4w15	DR5w11	DR6w19	DR7	DR8w2		
1283.21	AAYAAQGYKVLVLNPSVAATLGFEGAY	HCV NS3 1242-1267	4.5	350	5.2	567	143	5.1	89	288	34	175	9
1283.20	GYKVLVLNPSVAATL AQGYKVLVLNPSVAA	HCV NS3 1253 HCV NS3 1251	6.0	650	7.9	224	74	5.9	833	175	375	298	9
F98.03	AAAYAAQGYKVLVLNPSVAAT	HCV NS3 1242	2.9	48	483	18	124	103	11	96	60	240	9
F98.05	GYKVLVLNPSVAAT	HCV NS3 1248-1261	1.4	39	365	7.8	141	75	3.5	126	21	266	9
F98.04	GYKVLVLNPSVAATLGFEGAY GEGAAQWMNRLLFAASRGHVS	HCV NS3 1248-1267 HCV NS4 1914-1935	3.5	42	815	9.7	1500	240	4.1	23	80	20	8
1283.44	MNRLIAFASRGHVS	HCV NS4 1921	66	4.8	158	632	585	45	7.3	227	102	313	147
F134.08	GEGAAQWMNRLLFAASRGHNV	HCV NS4 1914	3.2	182	361	345	24	152	4.8	221	158	618	6
1283.16	SKGWRLLAAPTAYAQ	HCV NS3 1025	0.36	125	23	200000	667	417	745	19	962	54	1190
1283.55	GSSYGFQYSPGQRVE	HCV NS5 2641	11	16	625	625	200000	19	156	1900	68	384	8
1283.61	ASCLRKLGVPPLRVW	HCV NS5 2939	5.0	16	217	625	78	645	2500	862	671	882	7
F134.05	NFISGQYLAGLSTLPGNPA	HCV NS4 1772	10	606	84	29	29	29	29	70	441	441	6

Shading indicates IC50 > 1 μ M.
A dash (-) indicates IC50 > 20 μ M.

Table XXXV. HLA-DR binding capacity of 3 DR3 motif-containing peptides

Peptide	Sequence	Source	DR3 binding (IC50 nM)
35.0106	VVVVATDALMGTGTYG	HCV 1437	427
35.0107	TVDFSLDPPTFIETT	HCV 1466	235
1283,25	GRHLIFCHSKKKCDE	HCV NS3 1393	ND

Table XXXVIIa: HCV-derived CTL epitope candidates

Peptide	Molecule	1st Position	Sequence	Conserv.	Selection criteria
1073.05	NS4	1812	LLFNILGGWV	85	A2-superotype
1090.18	NS1/E2	728	FLLLADARV	92	A2-superotype
1013.02	NS4	1590	YLVAYQATV	85	A2-superotype
1090.22	NS5	2611	RLIVFPDLGV	79	A2-superotype
1013.1002	CORE	132	DLMGYIPLV	79	A2-superotype
24.0073	NS4	1920	WMNRLJFA	100	A2-superotype
24.0075	NS4	1666	VLVGGVLAA	85	A2-superotype
1174.08	NS4	1769	HMWNFISGI	92	A2-superotype
1073.06	NS4	1851	ILAGYGAGV	79	A2-superotype
1073.07	CORE	35	YLLPQQGPRL	92	A2-superotype
24.0071	NS1/E2	726	LLFLLLADA	100	A2-superotype
1.0119	LORF	1131	YLVTRHADV	85	A2-superotype
1.0952	CORE	51	KTSERSQPR	92	A3-superotype
1073.11	CORE	43	RLGVRATRK	79	A3-superotype
1.0955	ENV1	290	QLFTFSPPR	79	A3-superotype
1073.13	NS1/E2	632	RMYVGGVEHR	100	A3-superotype
1.0123	NS3	1396	LIFCHSKKK	100	A3-superotype
1073.10	NS4	1863	GVAGALVAFK	85	A3-superotype
24.0090	NS4	1864	VAGALVAFK	85	A3-superotype
24.0086	NS3	1262	TLGFGAYMSK	85	A3-superotype
F104.01	NS5	3003	VGIVYLLPNR	79	A31
1145.12	Core	169	LPGCSFSIF	92	B7-superotype
29.0035	NS3	1378	IPFYGKAI	92	B7-superotype
13.0019	NS5	2922	LSAFSLHSY	79	A1
1069.62	NS3	1128	CTCGSSSDLY	79	A1
24.0092	NS4	1765	FWAKHMWNF	85	A24

Table XXXVIIb: HCV-derived HTL epitope candidates

Region	Peptide	Motif ¹	Sequence
HCV NS3 1025-1039	1283.16	DR	SKGWRLLAPITAYAQ`
HCV NS3 1242-1267	F98.03	DR	AAYAAQGYKVVLVNPNSVAAT.
HCV NS3 1393-1407	1283.25	DR3	GRHLIFCHSKKKCDE`
HCV NS3 1437-1451	35.0106	DR3	VVVVATDALMTGTYTG`
HCV NS3 1466-1480	35.0107	DR3	TVDFSLDPTFTIETT`
HCV NS4 1772-1790	F134.05	DR	NFISGIQYLAGLSTLPGNPA`
HCV NS4 1914-1935	F134.08	DR	GEGAVQWMNRLLIAFASRGHNH`
HCV NSS 2641-2655	1283.55	DR	GSSYGFQYSRPGQERVE`
HCV NSS 2939-2953	1283.61	DR	ASCLRKLGVPPLRVW`

1. Peptides identified on the basis of either the DR P1-P6 supermotif or by use of the DR 1-4-7 algorithms are indicated by 'DR'. Peptides identified using the DR3 motif are indicated by 'DR3'.

Table XXXVII. Estimated population coverage by a panel of HCV derived HLAB epitopes

Antigen	Alleles	Representative assay	No. of epitopes ²	Population coverage (phenotypic frequency)					
				Cauc.	Blk.	Jpn.	Chn.	Hisp.	Avg.
DR1	DRB1*0101-03	DR1	6	18.5	8.4	10.7	4.5	10.1	10.4
DR2	DRB1*1501-03	DR2w2 β1	3	19.9	14.8	30.9	22.0	15.0	20.5
DR2	DRB5*0101	DR2w2 β2	6	-	-	-	-	-	-
DR3	DRB1*0301-2	DR3	2	17.7	19.5	0.40	7.3	14.4	11.9
DR4	DRB1*0401-12	DR4w4	5	23.6	6.1	40.4	21.9	29.8	24.4
DR4	DRB1*0401-12	DR4w15	3	-	-	-	-	-	-
DR7	DRB1*0701-02	DR7	5	26.2	11.1	1.0	15.0	16.6	14.0
DR8	DRB1*0801-5	DR8w2	5	5.5	10.9	25.0	10.7	23.3	15.1
DR9	DRB1*090111,09012	DR9	3	3.6	4.7	24.5	19.9	6.7	11.9
DR11	DRB1*1101-05	DR5w11	5	17.0	18.0	4.9	19.4	18.1	15.5
DR13	DRB1*1301-06	DR6w19	2	21.7	16.5	14.6	12.2	10.5	15.1
Total			98.5	95.1	97.1	91.3	94.3	95.1	

1. Total population coverage has been adjusted to account for the presence of DRX in many ethnic populations. It has been assumed that the range of specificities represented by DRX alleles will mirror those of previously characterized HLAB-DR alleles. The proportion of DRX incorporated under each motif is representative of the frequency of the motif in the remainder of the population. Total coverage has not been adjusted to account for unknown gene types.

2. Number of epitopes represents a minimal estimate, considering only the epitopes shown in Table 6. Additional alleles possibly bound by nested epitopes have not been accounted.

TABLE Ia

SUPERMOTIFS	POSITION	POSITION	POSITION
	2 (Primary Anchor)	3 (Primary Anchor)	C Terminus (Primary Anchor)
A1	T, I, L, V, M, S		F, W, Y
A2	V, Q, A, T		I, V, L, M, A, T
A3	V, S, M, A, T, L, I		R, K
A24	Y, F, W, I, V, L, M, T		F, I, Y, W, L, M
B7	P		V, I, L, F, M, W, Y, A
B27	R, H, K		F, Y, L, W, M, I, V, A
B58	A, T, S		F, W, Y, L, I, V, M, A
B62	Q, L, I, V, M, P		F, W, Y, M, I, V, L, A
<hr/>			
MOTIFS			
A1	T, S, M		Y
A1		D, E, A, S	Y
A2.1	V, Q, A, T*		V, L, I, M, A, T
A3.2	L, M, V, I, S, A, T, F, C, G, D		K, Y, R, H, F, A
A11	V, T, M, L, I, S, A, G, N, C, D, F		K, R, H, Y
A24	Y, F, W		F, L, I, W

*If 2 is V, or Q, the C-term is not L

Bolded residues are preferred, italicized residues are less preferred: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

WHAT IS CLAIMED IS

1. A composition comprising a prepared hepatitis C virus (HCV) epitope consisting of an amino acid sequence selected from the group consisting of :
FLLADARV, YLVAYQATV, RLIVFPDLGV,
DLMGYIPLV, WMNRLIAFA, VLVGGVLAA,
HMWNFISGI, ILAGYGAGV, YLLPRRGPRL,
LLFLLLADA, YLVTRHADV, KTSERSQPR,
RLGVRATRK, QLFTFSPRR, RMYVGGVEHR,
LIFCHSKKK, GVAGALVAFK, VAGALVAFK,
TLGFGAYMSK, LPGCSFSIF, LSAFSLHSY,
CTCGSSDLY, FWAKHMWNF, SKGWRLLAPITAYAQ,
AYAAQGYKVLVLNPSVAAT, GRHLIFCHSKKKCDE, VVVVATDALMTGYTG,
TVDFSLDPTFTIETT, NFISGIQYLAGLSTLPGNPA,
GEGAVQWMNRLIAFASRGNHV, GSSYGFQYSPGQRVE, ASCLRKLGVPLRVW,
and LTCGFADLMGY.
2. The composition of claim 1, further comprising two epitopes selected from the group in claim 1.
3. The composition of claim 2, further comprising three epitopes selected from the group in claim 1.
4. The composition of claim 1, wherein the composition further comprises a CTL epitope selected from the group consisting of LTDPHITA, LADGGCSEGGAY, RMILMTHFF, VMGSSYGF, FWAKHMWNFI, LLFNILGGWV, IPFYGKAI, and VGIYLLPNR.
5. The composition of claim 1, wherein the composition further comprises an HTL epitope.
6. The composition of claim 5, wherein the HTL epitope is a pan DR binding molecule.

7. The composition of claim 1, wherein the epitope is on or within a liposome.
8. The composition of claim 1, wherein the peptide is joined to a lipid.
9. The composition of claim 1, wherein the epitope is bound to an HLA heavy chain, β 2-microglobulin, and streptavidin complex, whereby a tetramer is formed.
10. The composition of claim 1, wherein the epitope is bound to an HLA molecule on an antigen presenting cell.
11. The composition of claim 10, wherein the antigen presenting cell is a dendritic cell.
12. The composition of claim 1, the composition further comprising a pharmaceutical excipient.
13. The composition of claim 1, further wherein the epitope is in a unit dose form.
14. A composition comprising a prepared peptide of less than 250 amino acid residues comprising at least two hepatitis C virus (HCV) peptide epitopes selected from the group consisting of:

FLLLADARV,	YLVAYQATV,	RLIVFPDLGV,
DLMGYIPLV,	WMNRLIAFA,	VLVGGVLAA,
HMWNFISGI,	ILAGYGAGV,	YLLPRRGPRL,
LLFLLLADA,	YLVTRHADV,	KTSERSQPR,
RLGVRATRK,	QLFTFSPRR,	RYMYVGGVEHR,
LIFCHSKKK,	GVAGALVAFK,	VAGALVAFK,
TLGFGAYMSK,	LPGCSFSIF,	LSAFSLHSY,
CTCGSSDLY,	FWAKHMWNF,	SKGWRLLAPITAYAQ,

AAYAAQGYKVLVLNPSVAAT, GRHLIFCHSKKKCDE, VVVVATDALMTGYTG,
TVDFSLDPTFTIETT, NFISGIQYLAGLSTLPGNPA,
GEGAVQWMNRLIAFASRGNHV, GSSYGFQYSPGQRVE, ASCLRKLGVPPRLRVW,
and LTCGFADLMGY.

15. The composition of claim 14, wherein at least two epitopes are linked via a spacer.

16. The composition of claim 14, further comprising a third epitope.

17. The composition of claim 16, wherein the third epitope is selected from the group consisting of LTDPHITA, LADGGCSGGAY, RMILMTHFF, VMGSSYGF, FWAKHMWNFI, LLFNILGGWV, IPFYGKAI, and VGIYLLPNR.

18. The composition of claim 16, further comprising a third epitope that is an HTL epitope.

19. The composition of claim 18, wherein the HTL epitope is a panDR binding molecule.

20. The composition of claim 14, wherein the peptide is on or within a liposome.

21. The composition of claim 14, wherein the peptide is joined to a lipid.

22. The composition of claim 14, wherein the peptide further comprises at least three of the epitopes in the group of claim 14.

23. The composition of claim 14, wherein the peptide further comprises at least four of the epitopes in the group of claim 14.

24. The composition of claim 14, wherein the peptide further comprises at least five of the epitopes in the group of claim 14.

25. The composition of claim 14, wherein the peptide further comprises at least six of the epitopes in the group of claim 14.

26. The composition of claim 14, the composition further comprising a pharmaceutical excipient.

27. The composition of claim 14, further wherein the epitope is in a unit dose form.

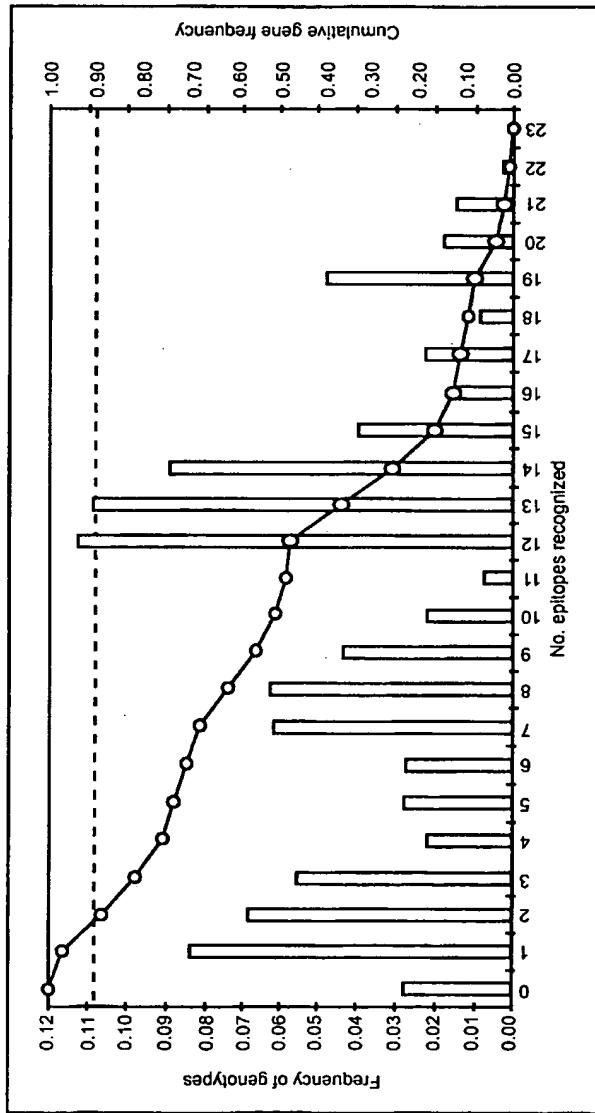
28. A composition comprising at least six prepared HCV epitopes each consisting of an amino acid sequence selected from the group consisting of:

FLLLADARV,	YLVAYQATV,	RLIVFPDLGV,
DLMGYIPLV,	WMNRLIAFA,	VLVGGVLAA,
HMWNFISGI,	ILAGYGAGV,	YLLPRRGPRL,
LLFLLLADA,	YLVTRHADV,	KTSERSQPR,
RLGVRATRK,	QLFTFSPRR,	RMYVGGVEHR,
LIFCHSKKK,	GVAGALVAFK,	VAGALVAFK,
TLGFGAYMSK,	LPGCSFSIF,	LSAFSLHSY,
CTCGSSDLY,	FWAKHMWNF,	SKGWRLLAPITAYAQ,
AAYAAQGYKVLVLNPSVAAT,	GRHLIFCHSKKKCDE,	VVVVATDALMTGYTG,
TVDFSLDPTFTIETT,	NFISGIQYLAGLSTLPGNPA,	
GEGAVQWMNRLIAFASRGNHV,	GSSYGFQYSPGQRVE,	ASCLRKLGVPPRLRVW,
and LTCGFADLMGY.		

29. The composition of claim 28, further comprising at least one epitope selected from the group consisting of LTDPHITA, LADGGCSGGAY, RMILMTHFF, VMGSSYGF, FWAKHMWNFI, LLFNILGGWV, IPFYGKAI, and VGIYLLPNR.

1 / 2

Monte Carlo population coverage analysis for
HCV candidate epitopes



Plot of total frequency of genotypes as a function of the number of HCV candidate epitopes bound by HLA-A and B alleles, in an average population. Genotype values were derived by averaging the gene frequencies in Caucasian, North American, Black, Japanese, Chinese, and Hispanic populations. Also shown is the cumulative frequency of genotypes.

Using currently available HLA typing data, a residual fraction (about 15%) of the genes, in an average population, are unspecified. To arrive at 100% accounting of genes, a fraction of the residual has been added for each hit population cluster in proportion to the relative frequency of the cluster within the HLA specified population. One peptide, 24.0086, was not incorporated into the present analysis.

FIG. 1

2 / 2

HVC Minigene

CTL Epitopes									
Kozak	SigSeq	Core 43	NS4 1590	NS3 1128	NS5 2611	Core 169	NS1/E2 632	NS4 1765	NS4 1863
A3	1073.11	1013.02	1069.62	1090.02	1145.12	1073.13	24.0092	1073.10	Core 132
		A2	A1	A2	B7	A3	A24	A3	A2
HTL Epitopes									
NS3 1253	NS4 1921	1437	NS5 2641	1466					
1283.21	1283.44	35.0106	1283.55	35.0107					
DR	DR	DR3	DR	DR3					

FIG. 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/19774

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : A61K 38/00, 38/04, 38/08, 38/10, 39/29, 39/295
US CL : 514/2,12,13,14,15, 885; 424/185.1, 189.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 514/2,12,13,14,15, 885; 424/185.1, 189.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
MEDLINE, BIOSIS, EMBASE, DERWENT WPI, WEST 2.0, search terms: author names, hcv, peptid?, HLA, htl, ctI.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim N
Y	WENTWORTH et al. Differences and similarities in the A2.1-restricted cytotoxic T cell repertoire in humans and human leukocyte antigen-transgenic mice. Eur. J. Immunol. 1996. Vol 26. pages 97-101, see entire document.	1-29
Y	US 5,736,142 A (SETTE et al.) 07 April 1998, see entire document.	1-29

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or prior date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means		
P document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
19 SEPTEMBER 2000	17 OCT 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer RON SCHWADRON Telephone No. (703) 308-0196